PCT





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

WO 97/18323 (51) International Patent Classification 6: (11) International Publication Number: **A2** C12Q 22 May 1997 (22.05.97) (43) International Publication Date: (81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, PCT/US96/19337 (21) International Application Number: CA. CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, 18 November 1996 (18.11.96) (22) International Filing Date: MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG). Eurasian (30) Priority Data: patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European US 16 November 1995 (16.11.95) 08/558.666 patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, US 27 February 1996 (27.02.96) 08/007.312 LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, 21 October 1996 (21.10.96) US 08/725,304 CM. GA. GN. ML, MR, NE, SN, TD, TG). (71) Applicant (for all designated States except US): ICOS CORPO-RATION [US/US]; 22021 20th Avenue, S.E., Bothell, WA Published Without international search report and to be republished 98021 (US). upon receipt of that report. (72) Inventors; and (75) Inventors/Applicants (for US only): HOEKSTRA, Merl, F. [US/US]; 10321 216th Street, S.E., Snohomish, WA 98290 (US). HOLTZMAN, Doug, A. [US/US]; 4308 4th Avenue, N.E., Seattle, WA 98105 (US). KEEGAN, Kathleen, S. [US/US]; 5812 Mercer Way, Mercer Island, WA 98040 (US). (74) Agent: NOLAND, Greta, E.; Marshall, O'Toole, Gerstein, Murray & Borun, 6300 Sears Tower, 233 South Wacker Drive, Chicago, IL 60606-6402 (US).

(54) Title: CELL CYCLE CHECKPOINT PIK-RELATED KINASE MATERIALS AND METHODS

(57) Abstract

The present invention generally relates to genes encoding cell cycle checkpoint phosphatidylinositol kinase (PIK)-related proteins essential to DNA damage responses in cells. These PIK-related kinases are required in regulatory pathways that arrest the cell cycle following DNA damage to allow DNA repair prior to mitosis or initiation of DNA replication. More particularly, the invention provides a novel human cell cycle checkpoint PIK-related kinase, MCCS1, and polynucleotide sequences encoding the MCSS1. Assays for identifying modulators of MCCS1 useful as, for example, chemotherapy and radiation adjuvants, are also provided by the invention. Further, assays for identifying modulators of the cell cycle checkpoint phosphatidylinositol kinase (PIK)-related protein identified as ATM are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
ΑT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
88	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	ΙE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO .	Romania
BY	Belarus	KG	Kyrgystan	RU.	Russian Federation
CA	Canada	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SG	
CH	Switzerland	KZ	Kazakhstan	SI	Singapore Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	
CM	Cameroon	LK	Sri Lanka	SN	Slovakia
CN	China	LR	Liberia	SZ	Senegal
CS	Czechoslovakia	LT	Lithuania		Swaziland
CZ	Czech Republic	LU	Luxembourg	TD	Chad
DE	Germany	LV	Latvia	TG	Togo
DK	Denmark	MC	Моласо	TJ	Tajikistan
EE	Estonia	MD		TT	Trinidad and Tobago
ES	Spain	MG	Republic of Moldova	UA	Ukraine
FI	Finland	ML	Madagascar Mali	UG	Uganda
FR	France			US	United States of America
GA	Gabon	MN	Mongolia	UZ	Uzbekistan
OA	Gaoon	MR	Mauritania	VN	Viet Nam

PCT/US96/19337

CELL CYCLE CHECKPOINT PIK-RELATED KINASE MATERIALS AND METHODS

FIELD OF THE INVENTION

The present invention generally relates to genes encoding cell-cycle checkpoint phosphatidylinositol kinase (PIK)-related genes and proteins essential to DNA damage responses in cells. The checkpoint kinases play a role in the surveillance of DNA damage that occurs as a result of replication errors, DNA mismatches, radiation treatment, or chemotherapeutic drugs. These kinases are required in regulatory pathways that lead to cell cycle arrest following DNA damage, giving the cell notice and time to correct lesions prior to the initiation of DNA replication. More particularly, the invention relates to a novel human PIK-related kinase, Mammalian Cell Cycle Surveillance 1 (MCCS1), polynucleotides encoding the PIK-related kinase, and methods for assaying and modulating the enzymatic activity of the kinase and related kinases.

15

20

25

5

10

BACKGROUND

The process of eukaryotic cell growth and division is the somatic or mitotic cell cycle which consists of four phases, the G_1 phase, the S phase, the G_2 phase, and the M phase. The G_1 , S, and G_2 phases are collectively referred to as interphase of the cell cycle. The cell cycle is structurally and functionally conserved in its basic process and mode of regulation across all eukaryotic species. During the G_1 (gap) phase, biosynthetic activities of the cell progress at a high rate. The S (synthesis) phase begins when DNA synthesis starts and ends when the DNA content of the nucleus of the cell has been replicated and two identical sets of chromosomes are formed. The cell then enters the G_2 (gap) phase which continues until mitosis starts. In mitosis, the chromosomes pair and separate and two new nuclei form, and in cytokinesis the cell itself splits into two daughter cells each receiving one nucleus

10

15

20

25

30

containing one of the two sets of chromosomes. Mitosis and cytokinesis together form the M (mitosis) phase of the cell cycle. Cytokinesis terminates the M phase and marks the beginning of interphase of the next cell cycle. The sequence in which the events in the cell cycle proceed is tightly regulated such that the initiation of one cell cycle event is dependent on the completion of the prior cell cycle event. This allows fidelity in the duplication and segregation of genetic material from one generation of cells to the next.

The term "cell cycle checkpoints" refers to the proteins, signals, processes, and feedback controls that integrate discontinuous events during cellular replication, in order to maintain essential dependencies within the cell cycle. The present invention specifically relates to the cell cycle checkpoint that ensures that mitosis is delayed until the completion of DNA synthesis and/or the accurate repair of DNA damage occurs.

Failure of cell cycle checkpoints predisposes individuals to or directly causes many disease states such as cancer, ataxia telangiectasia, embryo abnormalities, and various immunological defects associated with aberrant B and T cell development. The latter are associated with pathological states such as lupus, arthritis and autoimmune diseases. Intense research efforts have therefore focused on identifying cell cycle checkpoints and the proteins essential for the function of the checkpoints.

Genetic analysis in the yeasts Schizosaccharomyces pombe and Saccharomyces cerevisiae has identified a number of genes important for cell cycle arrest and DNA repair responses to ionizing radiation (IR). For a review, see Carr and Hoekstra, Trends in Cell Biology, 5: 32-40 (1995). One such gene, identified in both yeasts, is required for a DNA damage checkpoint which arrests the cell cycle at the G2 phase, as well as a related checkpoint which monitors the completion of DNA synthesis and arrests the cell cycle at the S phase. The gene is named rad3+ in S. pombe [Seaton et al., Gene, 119: 83-89 (1992)], MECI/ESR1 in S. cerevisiae [Kato et al., Nuc. Acids. Res., 22(15): 3104-3112 (1994)], and is hereinafter referred to as rad3+. Cells having mutations in rad3+ fail to either sense or appropriately respond to DNA damage and subsequently lose viability more rapidly than wild type cells after exposure to clastogenic agents or events (e.g., IR, DNA damaging agents,

and mutations affecting chromosomal integrity). See Weinert et al., GENES & DEVELOPMENT, 8: 652-665 (1994) and Al-Khodairy et al., EMBO J., 11(4): 1343-1350 (1992). This sensitivity to IR (radiosensitivity) can be caused by defects in checkpoint responses or defects in direct DNA repair reactions.

5

10

15

20

The product of the rad3 + gene is an approximately 270 kD protein that falls into a growing family of high molecular weight PIK-related kinases. Hunter, Cell, 83: 1-4 (1995) for a discussion of this family of kinases. The primary structures of the catalytic domains found in members of this gene family are closely related to well characterized phosphatidylinositol kinases. This structural relationship initially suggested that these PIK-related kinases might be capable of phosphorylating When the substrate specificity of the PIK-related kinases is examined, however, these enzymes appear to function as protein kinases and have yet to be demonstrated to phosphorylate phosphatidylinositides. Hartley et al., Cell, 82: 849856 (1995) reports that purified preparations highly active in protein kinase assays failed to show lipid kinase activity. Additional PIK-related kinases identified include: the TEL1 gene product from S. cerevisiae which affects telomere length [Greenwell et al., Cell, 82: 823-829 (1995)], and Mei41+ gene product from Drosophila melanogaster which is important for a G2 checkpoint and meiotic development [Hari et al., Cell, 82: 815-821 (1995)], the DNA-PK gene product from mouse which is important in immunoglobulin rearrangements and processing of DNA double strand breaks, and the FRAP gene product which is important in the G1/S transition [Brown, E. et al., Nature, 377:441-446 (1995)]. Mutations in the DNA-PK gene can result in the Severe Combined Immunodeficiency Syndrome (SCID) defect (Hartley et al., supra).

25

30

In humans, less is known about the molecular components required for checkpoint function. One component of the mammalian checkpoint machinery has been identified through the analysis of the human disease syndrome ataxiatelangiectasia (AT). Patients with AT show a diverse set of clinical symptoms, including predisposition to a variety of tumor types. Fibroblasts from AT patients are radiosensitive and fail to undergo cell cycle arrest following treatment with IR leading to a phenomenon termed radioresistant DNA synthesis. This is reminiscent of the S. pombe rad3 defect where cells fail to sense or respond appropriately to DNA damage.

10

15

20

Interestingly, the locus responsible for AT, the Ataxia-Telangiectasia Mutated (ATM) gene, was recently described in Savitsky et al., Science, 268: 1749-1753 (1995) and the partial cDNA encodes a protein with amino acid similarity to the rad3+ gene. Savitsky et al., Human Molecular Genetics, 4(11):2025-2032 (1995) describes isolation of a cDNA encoding full length ATM. The increased radiosensitivity of rad3+ yeast mutants and of mammalian cells lacking functional ATM suggests that these proteins may comprise a family of checkpoint proteins.

Kuerbitz et al., Proc. Natl. Acad. Sci. USA, 89: 7492-7495 (1992) establishes that the tumor suppressor p53 is required for a G1 checkpoint and cell cycle arrest observed following DNA damage. Irradiation of cells results in increased levels of p53 leading to the transcriptional activation of p53 responsive genes. One such p53-induced target is the product of the WAF1 gene (also called p21, CIP1, and sdil). WAF1 is a member of an expanding class of cell cycle regulators termed cyclin-dependent kinase inhibitory proteins. The activities of cyclin-dependent kinases control transit through the cell cycle. Transcriptional activation of WAF1 thus provides a direct link between DNA damage-dependent induction of p53 and the inhibition of kinases essential for cell cycle progression. See Elledge and Harper, Current Opinion in Cell Biology, 6: 847-852 (1994). Mutations in the p53 gene are one of the most common genetic alterations in human cancers. For example, Baker et al., Science, 244:217-221 (1989) reports that approximately 70% of human colorectal carcinomas contain deletions or mutant copies of the p53 gene. addition, Fearon et al., Cell, 61: 759-767 (1990) reports that breast, lung, bladder and brain tumors have been associated with loss of chromosome 17p, the region to which the p53 gene localizes.

25

30

At present there is relatively little known about the molecular components of the G2 checkpoints in mammalian cells. Caffeine is a chemical entity which abrogates G2 checkpoint control. Russell et al., Cancer Res., 55: 1639-1642 (1995) and Powell et al., Cancer Res., 55: 1643-1648 (1995) report that analysis of cell lines which differ only by the presence or absence of functional p53 demonstrated preferential caffeine-enhanced sensitization to IR in those cells lacking the p53-dependent G1 checkpoint. Thus, the conversion of potentially lethal damage into

lethal damage is greater in cells lacking the G1 and G2 checkpoints in comparison to cells containing an intact G1 checkpoint.

While certain cells undergo DNA damage-dependent cell cycle arrest, other cells appear to respond to DNA damage by initiating an intrinsic suicide program termed apoptosis or programmed cell death. The factors determining which process occurs are not fully understood. Recent work has demonstrated an important role for p53 both in the regulation of G1 cell cycle transitions and apoptosis. Symonds et al., Cell, 78: 703-711 (1994) describe p53-dependent apoptosis as suppressing tumor growth and progression in vivo.

10

15

5

High doses of radiation and chemotherapy are used to treat tumor cells in order to damage DNA so severely that the cells will die. However, even though tumor cells having mutations in the p53 gene are defective in a G1 checkpoint, they can still repair DNA damaged induced by radiation or chemotherapy. The present invention contemplates, for example, that inhibition of the G2 checkpoint in tumor cells should lead to a state in which tumor cells are incapable of repairing DNA damage therefore sensitizing the tumor cells to DNA damaging agents. Normal cells, containing intact G1 and G2 checkpoints, should still be able to repair DNA damage in the presence of a G2 checkpoint-specific inhibitor. Thus, treatment of tumors with a G2 checkpoint-specific inhibitor followed by radiation or chemotherapy should increase the efficacy of cell killing and thereby decrease the required doses of toxic DNA-damaging agents.

20

25

There thus exists a need in the art for identification of the mammalian proteins that are involved in the cell cycle checkpoints in order to develop therapies for the human disease states associated with defective cell cycle checkpoints and for the isolation of the genes encoding those proteins which in themselves may be useful as therapeutics or which would enable the development of therapeutically useful modulators of the proteins encoded by the genes.

SUMMARY OF THE INVENTION

The present invention provides novel human PIK-related kinases essential for a cell cycle checkpoint that responds in the G2 phase of the cell cycle to both damaged and unreplicated DNA.

10

15

20

25

30

In one of its aspects, the present invention provides purified and isolated polynucleotides (e.g., DNAs and RNAs, both coding and non-coding strands thereof) encoding the cell cycle checkpoint PIK-related kinase MCCS1 and polynucleotides encoding other cell cycle checkpoint PIK-related kinases that exhibit about 50, about 60, or about 65% nucleotide identity to the MCCS1 polynucleotide region encoding the MCCS1 kinase domain (MCCS1α nucleotides 6579 to 7562 of SEQ ID NO: 30 or MCCS1 β nucleotides 6457 to 7440 of SEQ ID NO: 32). Alternatively, the MCCS1-like PIK-related kinases exhibit about 40%, about 45%, or about 50% amino acid identity to the MCCS1 kinase domain (MCCS1 α amino acids 2083 to 2410 of SEQ ID NO: 31 or MCCS1 β amino acids 2152 to 2480 of SEQ ID NO: 33). Polynucleotides contemplated by the invention include genomic DNAs, RNAs, cDNAs and wholly or partially chemically synthesized DNAs. Preferred polynucleotides of the invention comprise the MCCS1 α DNA sequence set out in SEQ ID NO: 30, the partial MCCS1 β DNA sequence set out in SEQ ID NO: 3, the full length MCCS1\beta DNA sequence set out in SEQ ID NO: 32, and DNA sequences which hybridize to the noncoding strands thereof under stringent conditions or which would hybridize but for the redundancy of the genetic code. Exemplary stringent hybridization conditions are as follows: hybridization at 65°C in 3X SSC, 20mM NaPO₄ pH 6.8 and washing at 65°C in 0.2X SSC. It is understood by those of skill in the art that variation in these conditions occurs based on the length and GC nucleotide base content of the sequences to be hybridized. Formulas standard in the art are appropriate for determining exact hybridization conditions. See Sambrook et al., 9.47-9.51 in Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989). The MCCS1a DNA of SEQ ID NO: 30 was deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852, on November 3, 1995 as an insert in plasmid pBSHFB/HT2-27 in E. coli DH5α and was assigned ATCC Accession No. 69951. The MCCS1\beta DNA of SEQ ID NO: 32, was deposited with the ATCC on November 7, 1995 as an insert in plasmid 517 in E. coli DH5 α and was assigned ATCC Accession No. 69950.

The DNA sequence information provided by the present invention makes possible the identification and isolation of DNAs encoding related molecules

10

15

20

25

30

by well-known techniques such as DNA/DNA hybridization as described above and polymerase chain reaction (PCR) cloning. As one series of examples, knowledge of the sequence of a cDNA encoding MCCS1 makes possible the isolation by DNA/DNA hybridization of genomic DNA sequences encoding the kinase and expression control regulatory sequences such as promoters, operators and the like. Similarly, knowledge of a partial cDNA sequence encoding MCCS1 β make isolation of a complete cDNA possible. DNA/DNA hybridization procedures carried out with DNA sequences of the invention under stringent conditions are likewise expected to allow the isolation of DNAs encoding allelic variants of the PIK-related kinase; nonhuman species enzymes homologous to the PIK-related kinase; and other structurally related proteins sharing one or more of the enzymatic activities, or abilities to interact with members or regulators, of the cell cycle checkpoint pathway in which MCCS1 participates. Polynucleotides of the invention when detectably labelled are also useful in hybridization assays to detect the capacity of cells to synthesize MCCS1. The DNA sequence information provided by the present invention also makes possible the development, by homologous recombination or "knockout" strategies [see, Capecchi, Science, 244: 1288-1292 (1989)], of rodents that fail to express functional MCCS1 or that express a variant of MCCS1. Such rodents are useful as models for studying the activities of MCCS1 and MCCS1 modulators in vivo. Polynucleotides of the invention may also be the basis for diagnostic methods useful for identifying a genetic alteration(s) in the MCCS1 locus that underlies a disease state or states. Also made available by the invention are anti-sense polynucleotides relevant to regulating expression of MCCS1 by those cells which ordinarily express the same.

The invention also provides autonomously replicating recombinant constructions such as plasmid and viral DNA vectors incorporating polynucleotides of the invention, especially vectors in which the polynucleotides are functionally linked to an endogenous or heterologous expression control DNA sequence and a transcription terminator.

According to another aspect of the invention, host cells, especially unicellular host cells such as procaryotic and eukaryotic cells, are stably transformed or transfected with DNAs of the invention in a manner allowing expression of the PIK-related kinase therein. Host cells of the invention are conspicuously useful in

10

15

20

methods for the large scale production of MCCS1 wherein the cells are grown in a suitable culture medium and the desired enzymes are isolated from the cells or from the medium in which the cells are grown.

MCCS1 products having part or all of the amino acid sequence set out in SEQ ID NO: 31, SEQ ID NO: 4, or SEQ ID NO: 33 are contemplated. Use of mammalian host cells is expected to provide for such post-translational modifications (e.g., myristoylation, glycosylation, truncation, lipidation and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. The enzyme products of the invention may be full length polypeptides, fragments or variants. Variants comprise MCCS1 products wherein one or more of the specified (i.e., naturally encoded) amino acids is deleted or replaced or wherein one or more nonspecified amino acids are added: (1) without loss of the kinase activity specific to MCCS1; or (2) with disablement of the kinase activity specific to MCCS1; or (3) with disablement of the ability to interact with members or regulators of the cell cycle checkpoint pathway. Substrates of MCCS1 and proteins which interact with MCCS1 may be identified by various assays.

Substrates of MCCS1 may be identified by incorporating test compounds in assays for kinase activity. MCCS1 kinase is resuspended in kinase buffer and incubated either in the presence or absence of the test compound (e.g., casein, histone H1, or appropriate substrate peptide). Moles of phosphate transferred by the kinase to the test compound are measured by autoradiography or scintillation counting. Transfer of phosphate to the test compound is indicative that the test compound is a substrate of the kinase.

Interacting proteins may be identified by the following assays.

A first assay contemplated by the invention is a two-hybrid screen. The two-hybrid system was developed in yeast [Chien et al., Proc. Natl. Acad. Sci. USA, 88: 9578-9582 (1991)] and is based on functional in vivo reconstitution of a transcription factor which activates a reporter gene. Specifically, a polynucleotide encoding a protein that interacts with MCCS1 is isolated by: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA

25

30

10

15

20

25

30

binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding a first fusion of part or all of MCCS1 and either the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a library of second hybrid DNA sequences encoding second fusions of part or all of putative MCCS1 binding proteins and the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion; detecting binding of an MCCS1 interacting protein to MCCS1 in a particular host cell by detecting the production of reporter gene product in the host cell; and isolating second hybrid DNA sequences encoding the interacting protein from the particular host cell. Presently preferred for use in the assay are a lexA promoter to drive expression of the reporter gene, the lacZ reporter gene, a transcription factor comprising the lexA DNA binding domain and the GAL4 transactivation domain, and yeast host cells.

Other assays for identifying proteins that interact with MCCS1 may involve immobilizing MCCS1 or a test protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the amount of label bound. Bound label indicates that the test protein interacts with MCCS1.

Another type of assay for identifying MCCS1 interacting proteins involves immobilizing MCCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling a test protein with a compound capable of exciting the fluorescent agent, contacting the immobilized MCCS1 with the labelled test protein, detecting light emission by the fluorescent agent, and identifying interacting proteins as test proteins which result in the emission of light by the fluorescent agent. Alternatively, the putative interacting protein may be immobilized and MCCS1 may be labelled in the assay.

Also comprehended by the present invention are antibody products (e.g., monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, CDR-grafted antibodies and the like) and other binding proteins (such as those identified in the assays above) which are specific for the MCCS1 kinases of the invention. Binding proteins can be developed using isolated natural or recombinant enzymes. The binding proteins are useful, in turn, for purifying recombinant and

10

15

20

25

30

naturally occurring enzymes and identifying cells producing such enzymes. Specifically illistrating monoclonal antibodies of the invention are the monoclonal antibodies produced by hybridoma cell lines 224C and 224F which were deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive. Rockville, MD 20852 on November 7, 1996 and assigned ATCC Accession Nos. HB 12233 and HB 12234, respectively. Assays for the detection and quantification of proteins in cells and in fluids may involve a single antibody substance or multiple antibody substances in a "sandwich" assay format. The binding proteins are also manifestly useful in modulating (*i.e.*, blocking, inhibiting, or stimulating) enzyme/substrate or enzyme/regulator interactions. Anti-idiotypic antibodies specific for PIK-related kinase binding proteins are also contemplated.

The invention contemplates that mutations in the MCCS1 gene that result in loss of normal function of the MCCS1 gene product underlie human disease states in which failure of the G2 cell cycle checkpoint is involved. Gene therapy to restore MCCS1 activity would thus be indicated in treating those disease states (for example, testicular cancer). Delivery of a functional MCCS1 gene to appropriate cells is effected in vivo or ex vivo by use of viral vectors (e.g., adenovirus, adenoassociated virus, or a retrovirus) or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). For reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Alternatively, it is contemplated that in other human disease states preventing the expression of or inhibiting the activity of MCCS1 will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of MCCS1. Antisense nucleic acids (preferably 10 to 20 base pair oligonucleotides) capable of specifically binding to MCCS1 expression control sequences or MCCS1 RNA are introduced into cells (e.g., by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the MCCS1 target sequence in the cell and prevents transcription or translation of the target sequence. Phosphothioate and methylphosphate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense

10

15

20

25

30

oligonucleotides may be further modified by poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end.

Moreover, for example, if a particular form of cancer results from a mutation in a gene other than MCCS1 such as the p53 gene, an agent which inhibits the transcription or the enzymatic activity of MCCS1 and thus the G_2 cell cycle checkpoint may be used to render cancerous cells more sensitive to chemotherapy or radiation therapy. The therapeutic value of such an agent lies in the fact that current radiation therapy or chemotherapy in most cases does nothing to overcome the ability of the p53 mutant cancerous cell to sense and correct the DNA damage imposed as a result of the treatment. As a result, a cancer cell can simply repair the DNA damage. Modulating agents of the invention may therefore be chemotherapy and radiation adjuvants or may be directly active as chemotherapeutic drugs themselves.

Agents that modulate MCCS1 kinase activity may be identified by incubating a test compound with MCCS1 immunopurified from cells naturally expressing the PIK-related kinase, with MCCS1 obtained from recombinant procaryotic or eukaryotic host cells expressing the enzyme, or with purified MCCS1, and then determining the effect of the test compound on MCCS1 activity. The activity of the PIK-related kinase can be measured by determining the moles of ³²P-phosphate transferred by the kinase from gamma-³²P-ATP to either itself (autophosphorylation) or to an exogenous substrate such as a lipid or protein. The amount of phosphate incorporated into the substrate is measured by scintillation counting or autoradiography. An increase in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the test compound is an activator of said MCCS1 kinase. Conversely, a decrease in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the modulator is an inhibitor of said MCCS1 kinase. In another aspect, agents that modulate both MCCS1 and ATM or modulate one of the enzymes are also contemplated. Agents which modulate MCCS1 are screened in a kinase assay as described above in which ATM is the phosphorylating enzyme.

10

15

20

25

30

In a presently preferred assay, a MCCS1-specific antibody linked to agarose beads is incubated with a cell lysate prepared from host cells expressing the kinase. The beads are washed to remove proteins binding nonspecifically to the beads and the beads are then resuspended in kinase buffer. The reaction is initiated by the addition of gamma-³²P-ATP and an appropriate exogenous substrate such as lipid or peptide. The activity of the kinase is measured by determining the moles of ³²P-phosphate transferred either to the kinase itself or the added substrate. In a preferred embodiment the host cells lack endogenous MCCS1 and/or ATM kinase activity. The selectivity of a compound that modulates the kinase activity of MCCS1 can be evaluated by comparing its activity on MCCS1 to its activity on other known PIK-related kinases. The combination of the recombinant MCCS1 products of the invention with other recombinant PIK-related kinase products in a series of independent assays provides a system for developing selective modulators of MCCS1.

Furthermore, combinatorial libraries, peptide and peptide mimetics, defined chemical entities, oligonucleotides, and natural product libraries may be screened for activity as modulators in assays such as those described below.

For example, an assay for identifying modulators of MCCS1 kinase activity involves incubating an MCCS1 kinase preparation in kinase buffer with gamma-³²P-ATP and an exogenous kinase substrate, both in the presence and absence of a test compound, and measuring the moles of phosphate transferred to the substrate. An increase in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the test compound is an activator of said MCCS1 kinase. Conversely, a decrease in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the modulator is an inhibitor of said MCCS1 kinase.

Moreover, assays for identifying compounds that modulate interaction of MCCS1 with other proteins may involve: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA-binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding

ş.

a first fusion of part or all of MCCS1 and the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a second hybrid DNA sequence encoding part or all of a protein that interacts with MCCS1 and the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion; evaluating the effect of a test compound on the interaction between MCCS1 and the interacting protein by detecting binding of the interacting protein to MCCS1 in a particular host cell by measuring the production of reporter gene product in the host cell in the presence or absence of the test compound; and identifying modulating compounds as those test compounds altering production of the reported gene product in comparison to production of the reporter gene product in the absence of the modulating compound. Presently preferred for use in the assay are a lexA promoter to drive expression of the reporter gene, the lacZ reporter gene, a transcription factor comprising the lexA DNA binding domain and the GALA transactivation domain, and yeast host cells.

15

20

10

5

Another type of assay for identifying compounds that modulate the interaction between MCCS1 and an interacting protein involves immobilizing MCCS1 or a natural MCCS1 interacting protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the effect of a test compound on the amount of label bound wherein a reduction in the label bound in the present of the test compound compared to the amount of label bound in the absence of the test compound indicates that the test agent is an inhibitor of MCCS1 interaction with protein. Conversely, an increase in the bound in the presence of the test compound compared to the amount label bound in the absence of the compound indicates that the putative modulator is an activator of MCCS1 interaction with the protein.

25

30

Yet another method contemplated by the invention for identifying compounds that modulate the binding between MCCS1 and an interacting protein involves immobilizing MCCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling the interacting protein with a compound capable of exciting the fluorescent agent, contacting the immobilized MCCS1 with the labelled interacting protein in the presence and absence of a test compound, detecting light emission by the fluorescent agent, and identifying

modulating compounds as those test compounds that affect the emission of light by the fluorescent agent in comparison to the emission of light by the fluorescent agent in the absence of the test compound. Alternatively, the MCCS1 interacting protein may be immobilized and MCCS1 may be labelled in the assay.

5

The present invention further provides a cell-based complementation assay for identifying compounds which modulate the activity of MCCS1 or ATM. The assay involves complementation of a phenotypic trait associated with a genetic alteration in the cell. For example, the genetic alteration identified as esr1-1 results in cellular sensitivity to DNA damage in yeast cells [Kato et al., Nuc. Acids. Res., 22(15): 3104-3112 (1994)]. esr1-1 cells fail to either sense or appropriately response to DNA damage after exposure to DNA damaging agents such as ionizing radiation or clastogenic agents. The phenotypic trait of the genetically altered cell is complemented by transforming and expressing MCCS1 or ATM in the cell. The transformed cells are exposed to DNA damaging treatment (e.g. ionizing radiation) in the presence and absence of a test compound and sensitivity of the cells to DNA damage is measured. Agents that affect the cell sensitivity to DNA damaging activity of MCCS1 and/or ATM are identified as modulators.

15

20

10

Modulators of MCCS1 may affect its kinase activity, its localization in the cell, and/or its interaction with members of the cell cycle checkpoint pathway. MCCS1 modulators may be formulated in compositions comprising pharmaceutically acceptable carriers. Such compositions may additionally include chemotherapeutic agents. Dosage amounts indicated would be sufficient to result in modulation of MCCS1 activity in vivo. Selective modulators may include, for example, polypeptides or peptides which specifically bind to MCCS1 or MCCS1 nucleic acid, oligonucleotides which specifically bind to the PIK-related kinase or PIK-related kinase nucleic acid, and/or other non-peptide compounds (e.g., isolated or synthetic organic molecules) which specifically react with MCCS1 or MCCS1 nucleic acid. Mutant forms of MCCS1 which affect the enzymatic activity or cellular localization of wild-type MCCS1 are also contemplated by the invention. Presently preferred regions of the PIK-related kinases which are targets for the development of selective modulators include, for example, the following four regions: the MCCS1α amino terminal effector domain (amino acids 1 to 1081 of SEQ ID NO: 31), the MCCS1β

30

25

amino terminal effector domain (amino acids 1 to 1150 of SEQ ID NO: 33), the MCCS1 α rad3+ domain (amino acids 1082 to 2082 of SEQ ID NO: 31), the MCCS1 β rad3+ domain (amino acids 1151 to 2151 of SEQ ID NO: 33), the MCCS1 α PIK domain (amino acids 2083 to 2410 of SEQ ID NO: 31), and the MCCS1 β PIK domain (amino acids 2152 to 2480 of SEQ ID NO: 33).

10

15

20

25

30

DETAILED DESCRIPTION

The present invention is illustrated by the following examples. Example 1 details the isolation of cDNAs encoding MCCS1 kinases. Example 2 describes mapping of the human MCCS1 gene to human chromosome 3. recombinant expression of MCCS1 in E. coli and insect cells is respectively described in Examples 3 and 4. Example 4 also presents assays for measuring MCCS1 kinase activity. Example 5 describes the production of MCCS1-specific polyclonal and monoclonal antibodies. Example 6 reports the immunoprecipitation of MCCS1 kinase associated activity from mouse testes. Example 7 examines the expression of MCCS1 mRNA in various human tissues and cancer cell lines. Example 8 describes analyses of MCCS1 mRNA and protein expression in mouse testes. Example 9 describes analyses of MCCS1 protein expression in meiotic cells. Assays for substrates and interacting proteins of MCCS1 are described in Example 10. Example 11 describes modulators and assays for modulators of the kinase activity of MCCS1. Example 12 describes the cell-based complementation assay for identifying modulators of MCCS1 and/or ATM and Example 13 describes the kinase activity of ATM.

Example 1

cDNAs encoding the PIK-related kinase MCCS1 were isolated by a series of PCR reactions.

An alignment of the amino acid sequences of S. pombe rad3+ (Hari et al., supra) and S. cerevisiae MEC1 (Kato et al., supra) was the basis for design of seven degenerate oligonucleotides that encoded (or were complementary to) the regions of highest homology/lowest degeneracy between the sequences and contained

convenient restriction sites to facilitate cloning of amplification products. oligonucleotides were then used in a PCR-based assay to isolate a related human

sequence.

Initially, PCR amplifications were performed on cDNA preparations from rat T-cells, human peripheral blood mononuclear cells (PBMC), and S. cerevisiae genomic DNA. Five oligonucleotide pairs were used (oDH15a/oDH16, oDH15b/oDH16, oDH17a/oDH16, oDH15a/oDH17b, and oDH15b/oDH17b) for the primary amplifications. The sequences of the oligonucleotide primers included

inosines and are set out below in IUPAC nomenclature for degenerate nucleotide positions.

oDH15a (SEQ ID NO: 5)

5' GCA GAC GGA TCC GGI WCI GAY GGI AAY HTI TAY 3'

5 oDH15b (SEQ ID NO: 6)

5' GCA GAC GGA TCC GGI WCI GAY GGI AAY 3'

oDH16 (SEQ ID NO: 7)

5' GCA GAC GAA TTC RCA RTY RAA RTC IAC RTG 3'

oDH17a (SEQ ID NO: 8)

5' GCA GAC GGA TCC AAR TTY

CCI CCI RTI YTI TAY SAR TGG TT 3'

oDH17b (SEQ ID NO: 9)

5' GCA GAC GAA TCC AAC CAY

TSR TAI ARI AYI GGI GGR AAY TT 3'

PCR was performed on reaction mixtures of 1X PCR buffer (Perkin Elmer Cetus, Emeryville, California), 2-3μM oDH primers, 1.5mM MgCl₂, 200μM dNTPs, and 0.5 μl Amplitaq polymerase. The reaction was performed in a Perkin-Elmer Cetus Thermocycler Model 480 under the following conditions: denaturation at 94°C for 1 minute, annealing at 64°C for 2 minutes, and elongation at 72°C for 1 minute for 3 cycles. The procedure was then repeated using 60°C annealing temperature for 3 cycles, 56°C annealing for 3 cycles, and finished with denaturation at 94°C for 1 minute, annealing at 54°C (2 minutes, and elongation at 72°C for 1 minute for 30 cycles. PCR products were separated on 2 or 4% Tris Acetate EDTA (TAE) agarose gels, stained with ethidium bromide, and DNA products were visualized by UV fluorescence.

10

From the primary amplifications of yeast genomic DNA, rat T-cell cDNA, and human PBMC cDNA, only a single reaction with yeast genomic DNA (oDH17a/oDH16) gave a visible amplification product, resulting in a product that was the expected size for the region of the *S. cerevisiae MEC1* gene between these primers. Further analysis of the oDH17a/oDH16 amplifications that utilized rat T-cell and PBMC cDNA was therefore performed. To remove oligonucleotides and "primer dimers" that might interfere with subsequent PCR, primary reactions were purified prior to reamplification.

A "nested" PCR strategy was employed, and amplifications were repeated with primer pairs oDH18a/oDH16 and oDH18b/oDH16 under reaction conditions described above with cycle times of denaturation of 94°C for 1 minute, annealing at 55°C for 1 minute, and elongation at 72°C for 30 seconds for 30 cycles. The sequences of the oDH18a and oDH18b oligonucleotide primers included inosines and are set out below in IUPAC nomenclature for degenerate nucleotide positions.

15

20

5

10

oDH18a (SEQ ID NO: 10)

5' GCA GAC GGA TCC YTI GGI YTI GGI GAY CGI CA 3'

oDH18b (SEQ ID NO: 11)

5' GCA GAC GGA TCC YTI GGI YTI GGI GAY AGR CA 3'

An approximately 90 base pair (bp) product (the expected size amplification product for these primers) was seen in the reamplifications of the yeast genomic and human PBMC cDNA primary reactions. No 90 bp product was seen in the reamplification of the primary reaction on rat T-cell cDNA and this reaction was not analyzed further.

the oDH18a/oDH16 reaction yielded the appropriate size fragment during

In addition to the approximately 90 bp product, several other nonspecific bands were also present, though significantly fewer than were observed when the primary reactions were reamplified with oDH17a/oDH16. While the approximately 90 bp product was present in both the oDH18a/oDH16 and oDH18b/oDH16 reamplifications of the yeast genomic DNA primary reactions, only

10

15

20

25

30

reamplification of the human PMBC cDNA primary reaction. This was presumed to reflect codon usage in the human gene (compare primers oDH18a and oDH18b). The approximately 90 bp product from the oDH18a/oDH16 reamplification of the human PMBC cDNA primary reaction was gel purified and subcloned into the pBluescript SKII+ cloning vector (Stratagene, La Jolla, California) and sequenced.

Analysis of the sequence encoded by the 90 bp product indicated that the deduced amino acid sequence was similar to both S. cerevisiae MEC1 and S. pombe rad3+, but was not identical to either. To identify a larger region of coding sequence and extend the sequence comparison, a non-degenerate oligonucleotide, oDH23 5' GACGCAGAATTCACCAGTCAAAGAATCAAAGAG 3' (SEQ ID NO: 12), was synthesized for use in additional amplification reactions. Reamplification of the purified PBMC cDNA primary reaction with oDH17a/oDH23 led to the production of an amplification product of 174 bp. This fragment was then purified, subcloned and sequenced as described above. Computer analysis of the conceptual translation product confirmed its relationship (similar but not identical) to MEC1 and rad3+. This PCR fragment was then used as a probe to screen a plasmid library containing macrophage cDNA using the following hybridization conditions: incubation of nitrocellulose filters with radiolabelled probes in 3X SSC, 5X Denhardt's, 0.1% sarcosyl, 20mM NaPO₄ pH 6.8, 100 ug/ml single stranded salmon sperm DNA, for 18 to 24 hours at 65°C. Washes were done 3 times in 0.2X SSC, 0.1% SDS at 65°C for 30 minutes (with changes of wash buffer). Four positive clones were isolated, and the nucleotide sequence of each was determined. Computer analysis of the four sequences demonstrated that they were overlapping clones derived from a locus with homology to the rad3+ gene from S. pombe. Clone 517 (ATCC 69950) contained a 2.8 kbp insert and its DNA and deduced amino acid sequence are set out in SEQ ID NOs: 3 and 4, respectively. The clone contained an open reading frame encoding an amino terminal truncated protein product of 870 amino acids which were 39% identical to the COOH-terminus of rad3+. The protein product of the cDNA insert was named MCCS1 β .

The sequence of clone 517 was used to design the oligonucleotides, mo3 5'-CTACAGAGCCAAGGAG-3' (SEQ ID NO: 13) and mo6 5'-TCGAGCTATGCTACTAGTGGGC-3' (SEQ ID NO: 14), which were used to

10

15

20

25

30

generate a probe using a gel purified EcoRI fragment derived from clone 517 as a template. The PCR conditions were as follows: 50 ng DNA fragment. 1X PCR buffer (Perkin-Elmer Cetus), 1.5mM MgCl₂, 200 μ M dATP, dGTP, and TTP. 1 μ M dCTP, 50 μ Ci α^{32} P-dCTP, 10ng/ml each oligonucleotide, 1U AmpliTaq (Perkin-Elmer Cetus). The reaction was performed in a Perkin-Elmer Cetus Thermocycler Model 480 for an initial denaturing cycle at 94°C for 4 minutes followed by 20 cycles of 94°C for 15 seconds, 60°C for 15 seconds, 72°C for 30 seconds. Unincorporated nucleotides were removed using a Stratagene Nuc-trap Push Column.

Since Northern blot analyses showed that the expression of the mRNA corresponding to clone 517 was highest in testis, one million clones from a human testis cDNA library (Stratagene #939202) were screened with the PCR-generated probe and eleven clones were obtained. The two longest clones, HT2 and HT9, were chosen for analysis. HT2 contained a 4.7 Kb insert (corresponding to nucleotide 2974 of SEQ ID NO: 30 and extending further downstream than SEQ ID NO: 1) and HT9 contained a 5485 bp insert (corresponding to nucleotides 2152 to 7624 of SEQ ID NO: 30). Nucleotide sequence analysis revealed that in the region common to both cDNA clones there was a single base pair insertion of a T at nucleotide 3233 in HT9. This nucleotide insertion causes the predicted amino acid reading frame to shift and then terminate and is believed to be an error introduced by reverse transcriptase in clone HT9.

In order to isolate a clone containing an additional 2.5 Kb, one million clones from each of three additional cDNA libraries were screened: a human fetal brain cDNA library (Stratagene #93206), a human heart cDNA library (Stratagene #936207), and a human aorta cDNA library (Clontech Laboratories #HL1136a, Palo Alto, California). The sequence of the most 5' region of HT9 was utilized to design and synthesize two oligonucleotides, oHT9-1 5'-CCTAGTCCAGTAAAACTTGC-3' (SEQ ID NO: 15) and oHT9-4 5'-TTTGCGGCCCTTCCAATATC-3' (SEQ ID NO: 16) which were used to generate a 317 bp PCR probe under conditions described for generating the probe above. While no positive clones were isolated from the heart or aorta cDNA libraries, two positive clones were obtained from the fetal brain library. One of these clones, HFB2, included a cDNA 4.5 Kb insert which included

10

15

20

25

30

approximately 2300 bp of additional sequence. The HFB2 insert corresponds to nucleotides 1 to 3194 of SEQ ID NO: 30.

A composite cDNA encoding MCCS1 α was constructed from clones HFB2, HT9 and HT2. The three clones were joined together by digesting HFB2 with the restriction enzymes KpnI and Sall to generate a fragment to comprise the 5' end of the composite clone, digesting HT9 with KpnI and NotI to generate a fragment to comprise the 3' end of the composite clone, and then ligating isolated fragments to the vector pBS SK' (Stratagene) that had been digested with SalI and NotI. region of the HT9 fragment containing the one nucleotide insertion was replaced with an EcoRV fragment containing nucleotides 3174 to 5282 of clone HT2. The final plasmid containing a 7621 bp insert was named pBSHFB2HT2-27 (ATCC 69951). The DNA and deduced amino acid sequence of the insert are presented in SEQ ID NOs: 1 and 2, respectively. The coding domain of the cDNA initiates with an ATG at nucleotide 333 and ends with a termination codon at nucleotide 7560 predicting a coding sequence of 2409 amino acids and protein of 265 kD. The protein product of the cDNA insert was named MCCS1a. Subsequent sequence analysis of the insert in plasmid pBSHFB2HT2-27 (ATCC 69951) revealed sequencing errors in SEQ ID NO: 1. Corrected DNA and deduced amino acid sequences of the insert are set out in SEO ID NOs: 23 and 24, respectively. Even further sequence analysis of insert in plasmid pBSHFB2HT2-27 revealed sequencing error in SEQ ID NO: 23. At nucleotide position 6317 (SEQ ID NO: 23) a "G" was erroneously included and between positions 6338 and 6339 the sequence was missing an "A". The corrected sequences of MCCS1 α are provided in SEQ ID NOs: 30 and 31.

Comparison of the predicted amino acid sequence of MCCS1 α with the partial amino acid sequence of MCCS1 β predicted from clone 517 revealed the presence of a seventy amino acid deletion in the MCCS1 α product. The MCCS1 β clone 517 amino acid sequence corresponds to MCCS1 α amino acids 1611 to 2410 of SEQ ID NO: 31. The seventy amino acid deletion in MCCS1 α (i.e., where the seventy amino acids would be inserted to generate a product identical to MCCS1 β) occurs between amino acids 2065 and 2066 in SEQ ID NO: 31, seventeen amino acids upstream from the kinase domain. Since both clones maintain an open reading frame, cDNA clone pBSHFB2HT2-27 was apparently generated from alternatively

10

15

20

25

30

spliced mRNA. The carboxyl terminal domains containing the kinase domains are identical in MCCS1 α (amino acids 2083 to 2410 of SEQ ID NO: 31) and MCCS1 β clone (amino acids 543 to 870 of SEQ ID NO: 4).

A composite clone containing the complete coding sequence of MCCS1 β (with the seventy amino acid insert) is presented in SEQ ID NO: 32. The amino acid sequence deduced from the clone is presented in SEQ ID NO: 33. This clone is constructed by replacing the sequence between the BSTXI site, which cleaves after nucleotide 3229, and the NotI site in the polylinker sequence at the 3' end of pBSHFB2HT2-27 (SEQ ID NO: 1) with the sequence contained in HT2 between the BstXI site and the NotI site at the 3' end of HT2. Thus this clone contains sequences that are identical to MCCS1 α nucleotides 1 to 5159 of SEQ ID NO: 1 (encoding amino acids 1 to 1609 of SEQ ID NO: 2) linked to sequences that are identical to clone 517 nucleotides 1 to 2610 of SEQ ID NO: 3 (encoding amino acids 1 to 870 of SEQ ID NO: 4). As noted above, subsequent sequence analysis revealed errors in nucleotides 1 to 5159 of SEQ ID NO: 1. Corrected MCCS1\$\beta\$ DNA and deduced amino acid sequences that include the same corrections that appear in $MCCS1\alpha$ SEQ ID NOs: 23 and 24 are set out in SEQ ID NOs: 25 and 26. The SEQ ID NO: 25 clone represents a cDNA encoding a full length MCCS1 β kinase. Further sequences for MCCS1 β including corrections of errors identified in resequencing the MCCS1 α clone are presented in SEQ ID NOs: 32 and 33.

The MCCS1 products can be divided into three regions based on similarity to other PIK-related kinases: an amino terminal domain (MCCS1 α amino acids 1 to 1081 of SEQ ID NO: 31 and MCCS1 β amino acids 1 to 1150 of SEQ ID NO: 33), a region with similarity to rad3+ (MCCS1 α amino acids 1082 to 2082 of SEQ ID NO: 31 and MCSS1 β amino acids 1151 to 2151 of SEQ ID NO: 33) and a PIK domain (MCCS1 α amino acids 2083 to 2410 of SEQ ID NO: 31 and MCCS1 β amino acids 2152 to 2480 of SEQ ID NO: 33) including a kinase domain. The amino terminal region and rad3+ region are regulatory domains that modulate the kinase activity of the enzyme and are involved in interactions with associated proteins.

Results of comparisons of the nucleotide and amino acid sequence of MCCS1 α and MCCS1 β to the sequences of other PIK-related and non-PIK-related kinases are shown in Table 1. Specifically, the 3' end of MCCS1 α (nucleotides 6579)

10

15

20

to 7562 of SEQ ID NO: 30 encoding the kinase domain), the 3' end of MCCS1 β (nucleotides 1627 to 2379 of SEQ ID NO: 32 encoding the kinase domain), the rad3 + domain of MCCS1 α (nucleotides 3576 to 6578 of SEQ ID NO: 30), and the rad3+ domain of MCCS1 β (clone 517 nucleotides 1 to 1626 of SEQ ID NO: 3) were compared to the analogous region in human ATM [Savitsky et al., supra], human DNA-PK [Huntley et al., Cell, 82: 849-856 (1995)], human FRAP [Brown et al., supra], human p110 [Hu et al., Mol. Cell. Biol., 13(12): 7677-7688 (1993)], S. cerevisiae MEC1 [Weinert et al., Genes Dev., 8(6): 652-665 (1994), S. pombe rad3+ [Seaton et al., supra and Hari et al., Cell, 82: 815-821 (1995)] and an cAMPdependent protein kinase (PKA) [Beebe et al., Mol. Endocrinol., 4(3): 465-475 (1990)]. Percent identity of nucleotides is shown in the top line, percent identity of amino acids is shown in the middle line, and percent similarity of amino acids (i.e., including identical amino acids and conservative variations in amino acids) is shown in the bottom line for each kinase in Table 1. Conservative variation as used herein denotes biologically similar residues. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. In the Table, "ND" indicates a value was not determined either because the nucleotide sequence encoding the kinase (i.e.. rad3+) was not publically available or because the kinase (i.e., FRAP, p110 β , or PKA) lacks the particular domain being compared.

Table 1

	Protein Kinase	MCCS1 α /MCCS1 β Kinase Domain	MCCS1α rad3+ Domain	MCCS1 β rad3+ Domain
	S. pombe rad3+	ND	ND	ND
		56	22	30
		72	46	53
~	S. cerevisiae MECl	51	42	44
5		45	21	24
		63	46	49
	Human ATM	50	41	41
		38	22	24
		60	46	47
	Human DNA-PK	43	39	43
		29	19	20
		53	45	49
	Human FRAP	45	ND	ND
		37	ND	ND
		61	ND	ND
	Human p $110eta$	45	ND	ND
		24	ND	ND
		54	ND	ND
10	Human PKA	39	ND	ND
		16	ND	ND
		39	ND	ND

10

15

20

Example 2

The MCCS1 gene was mapped to chromosome 3 by a PCR-based assay. Human/rodent somatic cell hybrids containing various human chromosome panels available from the NIGMS Human Genetic Mutant Cell Repository [Drwinga et al., Genomics, 16: 311-314 (1993)] were used as templates.

Two oligonucleotide primers oDH23 (SEQ ID NO: 12) and oDH26 5' TGGTTTCTGAGAACATTCCCTGA 3' (SEQ ID NO: 19) based on the MCCS1 α cDNA sequence were utilized to amplify a portion of the gene. The primers generate 237 bp PCR products. PCR conditions consisted of 50 ng genomic DNA, 0.5 μ g of each primer, 200 μ M dNTPs, 1.5mM MgCl₂, 1X PCR buffer (Perkin Elmer-Cetus), and 1 unit of Amplitaq polymerase (Perkin-Elmer Cetus) in a 25 μ l reaction volume. The samples were denatured for 4 minutes and then cycled 35 times with denaturing, annealing, and extension times of 45 seconds, 30 seconds, and 45 seconds, respectively, in a Model 480 Cetus Thermocycler. Five μ l of the resulting PCR product was electrophoresed on a 3% agarose gel and stained with ethidium bromide. DNA corresponding to the human/rodent chromosome 3 hybrid yielded a positive amplification product.

In a second set of amplification reactions, the same oligonucleotide primers were used to sublocalize the MCCS1 gene to a specific region on chromosome 3. The templates for these amplifications consisted of DNA samples from patients with chromosome 3 truncations [Leach et al., Genomics, 24: 549-556 (1994)]. Amplifications were performed as described in the foregoing paragraph. The pattern of positive amplification products narrowed the localization to the interval between q21 and q25.1.

25

30

Example 3

Polynucleotides encoding carboxyl terminal portions of the PIK-related kinase MCCS1 β were expressed by recombinant techniques in $E.\ coli$.

Two *E. coli* expression plasmids were constructed that expressed either the COOH-terminal 423 or 571 amino acid residues of the kinase in the Pinpoint fusion protein expression/purification system (Promega, Madison, Wisconsin). Briefly, DNA sequences encoding the COOH-terminal portion of the kinase

(nucleotides 1339 to 2630 or nucleotides 898 to 2630 of SEQ ID NO: 3) were fused in frame to the COOH-terminus of a 13 kD peptide derived from the transcarboxylase complex from propionibacterium shermanii. This region undergoes biotination in *E. coli*, and thus provides a means for monitoring expression and purification of the fusion proteins. Expression was driven from the tac promoter in pinpoint Xa3. Fusion protein expression was induced with 0.1mM IPTG and confirmed using streptavidin alkaline phosphatase in a pseudo-Western format as described by the manufacturer.

Example 4

10

5

Recombinant versions of MCCS1 may also expressed in yeast or in SF9 insect cells using a baculovirus expression system. The FRAP kinase has been expressed, purified and is enzymatically active after expression in the baculovirus system [Brown et al., supra].

20

25

is used for protein purification.

heterologous peptide sequence, such as the FLAG tag MDYKDDDK (SEQ ID NO: 20) or a six-histidine tag, and reconstructed into the appropriate vectors. Once expressed in insect cells, a monoclonal antibody that recognizes the FLAG tag (Eastman Kodak, Rochester, New York) is used to purify large quantities of the FLAG-PIK-related kinase fusion protein. Infected insect cells are incubated for 48 hours and lysed in lysis buffer (25mM 2-glycerolphosphate, 50mM sodium phosphate pH 7.2, 0.5% Triton-X 100, 2mM EDTA, 2mM EGTA, 25 mM sodium fluoride, 100μM sodium vanadate, 1mM PMSF, 1μg/ml leupeptin, 1μg/ml pepstatin. 1mM benzamidine, and 2mM DTT). Expressed FLAG fusion proteins are purified over a column containing anti-FLAG antibody M2 affinity resin (Eastman Kodak). The column is washed with 20 column volumes of lysis buffer, then 5 column volumes of 0.5M lithium chloride, 50mM Tris pH 7.6, 1mM DTT, and then eluted either with 0.1M glycine pH 3.0 followed by immediate neutralization or by competitive elution with the FLAG peptide. For six-histidine tagged proteins, Ni-NTA agarose (Qiagen)

10

15

20

25

30

Shortly after the filing of parent application U.S.S.N. 08/558,666, a gene identified as ATR was described by Antony M. Carr and co-workers (personal communications). ATR appears to encode the same or a closely related protein to MCCS1 based on a comparison of amino acid sequences between ATR and MCCS1. The DNA and deduced amino acid sequences of ATR are presented in SEQ ID NOs: 28 and 29, respectively. The sequence differences between ATR and MCCS1 β are as follows. ATR includes an additional 98 amino acid residues at the N-terminus. At nucleotide position 1284 (SEQ ID NO: 32) there is a conservative base change from "A" in MCCS1 β to "T" in ATR and at nucleotide position 4176, there is an additional conservative base change from "C" in MCCS1 β to "T" in ATR.

The FLAG tag was fused at the amino-terminus of a truncated ATR molecule which lacked the first sixty-six ATR amino acids. The FLAG tag was oligos FLAG-ATR added bν PCR a s follows. The (5'-CGGGATCCGCCATGGACTACAAGGACGATGACAAGATGTTGCTTGATTTC-3'). And HFB24 (5'CTTAAGCCGCATGAGCACACCGTC-3') were used in the following PCR reaction: 100ng of pcDNAATR (obtained from Antony M. Carr) as template; 1X PCR buffer (Perkin-Elmer Cetus); 1.5 mM MgCl₂, 200µM each of dATP, dGTP, dCTP, and TTP, 10 ng/µl of each primer; 1U AmpliTaq (Perkin-Elmer Cetus). The reaction was denatured at 94°C for 4 minutes followed by 30 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 30 seconds. The resulting approximately 800 bp PCR product was digested with BamHI and NheI and was ligated to the 10kb fragment of the mammalian ATR expression plasmid, pcDNAATR digested with BamHI and BstXI along with the remainder of the ATR coding sequences contained on a 2.5 kb BstXI to NheI fragment. Sequence analysis confirmed the addition of the FLAG tag. The insert contained within this plasmid was then used to construct a baculovirus expression plasmid that would express the FLAG tagged ATR truncate. The 5' end of ATR contained on a BamHI to BstXI fragment and the 3' end of ATR contained on a BstXI to SalI fragment derived from pBTM ATR were ligated to the baculovirus expression vector, pFB (Gibco/BRL) that had been digested with BamHI and Sall. This plasmid was designated pFMBCCSβFLAG.

10

15

20

25

30

The full coding region of ATR was fused at the amino terminus to the six histidine tag b y PCR. Oligonucleotides (5'-CGGGATCCAGCATCACCATCACCATCACATGGGGGAACATGGGC-3') and Frp1R ("5'-CATGACCACTGGCCATTCCACACG-3') were used in a PCR reaction to add the six histidine tag to sequences encoding the amino-terminus of ATR. PCR conditions were as follows: 100 ng of PstA 12ATR (obtained from Antony M. Carr) was used as template; 1X PCR buffer (Perkin-Elmer Cetus); 1.5 mM MgCl₂, 200μM each of dATP, dGTP, dCTP, and TTP, 10 ng/μl of each primer; 1U AmpliTaq (Perkine-Elmer Cetus). The reaction was denatured at 94°C for 4 minutes followed by 25 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 30 seconds. The approximately 800 bp PCR product was digested with BamHI and MscI and ligated to two other fragments: a 10kb fragment from pcDNAATR digested with BamHI and BstXI and an approximately 3 kb MscI to BstXI fragment containing the remainder of the ATR coding sequence. The addition of the six histidine tag was verified by sequence analysis. The resulting plasmid encoding a six-histidine tagged full length ATR molecule was designated pcDNA6his ATR.

To construct a baculovirus expression plasmid that expressed the entire coding sequence of ATR, the 1.2 kb BamHI to AgeI fragment from pFBMCCS β FLAG was ligated to the BamHI to AgeI fragment from pcDNA6his ATR. The resulting plasmid, designated pFB/HisX6MCCS-1 plasmid was transformed into the E.coli strain, DH5 α (Gibco/BRL) for screening of recombinants. This plasmid was purified by using the Promega "Wizard" mini-prep kit, then transformed into E. coli α SF9 cells (Invitrogen) using the Cellfectin protocol described by Gibco/BRL.

Forty eight hours after transfection, the SF9 cell pellet and baculovirus produced by the transfected cells were harvested. The virus was stored at 4°C in Grace's Complete media containing 10% FBS, Pennicillin-Streptomycin, and Gentamicin. This viral prep was used to make a high titer (P2) virus stock. The P2 virus stock was used to infect a 50 ml culture of SF9 cells. The cells were collected 48 hours after infection and centrifuged at low speed to pellet the cells without lysis. The cell pellet was stored at -20°C for 24 hours before lysis. The cells were lysed

WO 97/18323 PCT/US96/19337

-29-

in 5 ml of lysis buffer (50 mM Tris, pH 8.0; 500 mM NaCl; 1% NP40; 100 μ m PMSF). Expression of ATR was confirmed by immunoblot using the polyclonal antibody anti-AgDH2 as a probe. The FBHisX6 ATR baculovirus produced an approximately 300 kDa protein that was immunoreactive with anti-AgDH2 antibodies and comigrated with a protein in a mouse testes cell extract.

The P2 virus stock was also used to infect a 2 liter culture of SF9 cells. The cells were collected 48 hours after infection, centrifuged at low speed to pellet the cells without lysis and stored at -20°C. A cell pellet from 150 mls of this culture was lysed in 7.5 ml of lysis buffer (50mM NaPO, pH7.2; 0.5% NP-40; 10mM imidazole, 25mM NaF, 100μM Na₃VO₄; 0.5mM AEBSF; 1 μg/ml leupeptin; 1μg/ml pepstatin A) and incubated on ice for 15 minutes. The lysate was then centrifuged for 30 minutes at 10,000 x g. The supernatant was removed and any DNA in the lysate resulting from broken nuclei was sheared by aspirating through an 20 gauge needle. Particulate matter was then removed by filtering through a 0.8 micron filter followed by a 0.2 micron filter. This cleared lysate was adjusted to contain 5 mM β -mercaptoethanol and 0.4 M NaC1. A 1 ml Ni-NTA-agarose column (Qiagen) was equilibrated in Buffer A (0.4 M NaC1; 5 mM β-mercaptoethanol; 0.1% Triton X-100; 50 mM NaPO₄ 10 mM imidazole; 25 mM NaF, 100 μ M Na₃VO₄; 0.5 mM AEBSF; 1 μ g/ml leupeptin; 1 μ g/ml pepstatin A) prior to loading the cleared lysate. The sample was loaded at a flow rate of 0.25 ml/minute, washed 5 ml of Buffer A and then eluted in 10 ml of a gradient of 50 to 500 mM imidazole in Buffer A. One half ml fractions were collected and was assayed for kinase activity as follows. Five μl of each fraction was incubated in kinase buffer, 10 μCi ³²PγATP, 10 μM ATP, and 5 μ g of substrate PHAS-1 (Stratagene) and incubated at 37°C for 20 minutes. The reaction was then spotted onto phosphocellulose spin columns and centrifuged at 2500x g, washed twice with 0.5 ml of 75 mM phosphoric acid and once with 0.5 ml absolute ethanol. The phosphocellulose disks were then transferred to scintillation vials and the counts per minutes incorporated into the PHAS-1 proteins were recorded. Fractions 4 through 9 were found to contain activity toward PHAS-1 and immunoblot analysis confirmed that ATR was also present in the same fractions.

5

10

15

20

25

30

MCCS1 encoding plasmid DNA was transformed into an esr1-1 diploid yeast strain (Mata leu2-1 his4-4 can1 ura3 cyh2 ade6 ade2 esr1-1/MAT a leu2-27 his4 trp1 met2 ade2 esr1-1), and cells were grown to mid-log phase in either galactose or glucose containing medium. Cells were pelleted, washed and all steps performed at 4° C. Cell pastes were resuspended in buffer (20 mM Tris at pH 8.0, 300 mM NaCl, 10% glycerol, 0.1 mM PMSF, 0.25 mg/ml pepstatin, leupeptin, and aprotinin) and lysed in a French Press or using glass beads. Lysis was verified by microscopy following a low-speed (10K) spin and a high-speed spin (100K), and the supernatant was loaded onto a 1.5 ml Ni-NTA agarose (Qiagen, Inc., Chatsworth, CA) column prewashed in 1x buffer. The column was washed with six column volumes of buffer. The column was eluted stepwise with 8 ml of 10 mM, 50 mM, 100 mM, and 250 mM imidazole in buffer. Fractions were collected and Western analysis was performed using 15 μ l of each elution peak. Kinase activity was measured as described above.

15

20

25

30

5

10

Example 5

Polyclonal and monoclonal antibodies specific for MCCS1 were generated by standard techniques in the art.

Two different bacterial expression plasmids, pGEX1-MEC and pGEX3-MEC, were constructed for the recombinant production of portions of the MCCS1 polypeptide as fusions to the COOH-terminus of glutathione S-transferase (GST). Both plasmids were used for the generation of antigens AgDH-2 and AgDH-3, from pGEX1-MEC and pGEX3-MEC respectively for use in a standard immunization protocol. pGEX1-MEC contains an EcoRI fragment encoding amino acid residues 566 to 870 of SEQ ID NO: 4 fused to GST in the pGEX1 vector (Pharmacia Biotech, Milwaukee, Wisconsin); pGEX3-MEC contains an Eco RI fragment encoding amino acid residues 118 to 567 of SEQ ID NO: 4 fused to GST in the pGEX3 vector (Pharmacia Biotech). Induction of the pGEX tac promoter with 0.1mM IPTG led to high level expression of each fusion protein in an insoluble form (inclusion bodies). Following lysis of induced cultures with a French pressure cell, AgDH-2 and AgDH-3 extracts were centrifuged through a 35% sucrose solution containing 0.1M NaCl,

WO 97/18323 PCT/US96/19337

-31-

0.01M Tris pH7.5, and 0.001M EDTA (STE). Pellets were then washed twice and resuspended in STE.

For the generation of polyclonal antisera in rabbits, AgDH-2 and AgDH-3 were further purified using preparative SDS polyacrylamide gel electrophoresis and electroelution of each antigen from gel slices. Primary immunization of female New Zeaiand White rabbits was with 200 μ g of each antigen mixed with complete Freund's adjuvant injected at multiple sites subcutaneously. Subsequent immunizations were with 100 μ g antigen mixed with incomplete Freund's adjuvant at approximately 21 day intervals, and test bleeds were taken after immunizations 3, 4 and 5. Western blot analysis of extracts of human testis tissue demonstrates antibody reactivity against an approximately 270 kD protein in immune but not preimmune antisera. In addition, the immune sera showed reactivity against the MCCS1 pinpoint fusion proteins described in Example 3, providing evidence of the generation of MCCS1-specific antibodies.

15

20

25

30

10

5

The MCCS1-specific antibodies were purified as follows. Inclusion body preparations of AgDH-2 and AgDH-3 were coupled to cyanogen bromide (CNBr)-activated Sepharose (Pharmacia, Alameda, CA). Two mg of antigen were solubilized in 1% SDS (4.5 ml final volume) and dialyzed overnight against Coupling Buffer (0.1M NaHCO₃/0.1% SDS). 0.5 ml of 5M NaCl were added to each antigen preparation prior to incubation with the CNBr Sepharose. 0.4 gm of freeze-dried CNBr Sepharose (per antigen) were resuspended in 1 mM HCl and washed in a scintered glass funnel with 250 ml 1 mM HCl added in several aliquots over 15 minutes. The HCl-washed CNBr Sepharose was then removed to a 15 ml snap cap tube and washed twice with 5 ml of Coupling Buffer. Dialyzed antigen preps were added to the washed Sepharose and then incubated at room temperature for 1.5 hours on a slowly rotating wheel. The Sepharose was washed once with 5 ml of Coupling Buffer, once with 10 ml of 0.1M Tris pH8.0, and then incubated in 10 ml 0.1M Tris 8.0 for 2 hours at room temperature to block any remaining reactive groups on the resin. Coupling efficiency was 60-80% as judged by SDS-PAGE analysis. The antigen columns were then washed with 15 ml of 6M Guanidine HCl (to remove uncoupled antigen), 25 ml of Buffer A (50mM Tris pH 7.4), 15 ml of Buffer B (4.5M MgCl₃/lmg/ml BSA/50mM Tris 7.4), and then 50 ml of Buffer A. Thirty ml

of rabbit serum from immunized animals (rabbit 4747 immunized with AgDH-3 and rabbit 4779 immunized with AgDH-2) were passed over the appropriate antigen column over the course of 3 hours. The columns were then washed with 20 ml of Buffer A, 40 ml of 1M Guanidine HCl, and then equilibrated with an additional 20 ml of Buffer A. Anti-AgDH-3 or Anti-AgDH-2 antibodies were then eluted off the antigen columns with 10 ml of Buffer B. One ml fractions were collected, IgG-containing fractions were pooled and dialyzed against 1 L of phosphate buffered saline (PBS) for 3 hours, and then overnight against 1 L of PBS containing 35% glycerol.

10

15

5

Antipeptide antibodies were generated against the human ATM protein by coupling a 15-amino-acid peptide (residues 1359-1373) to Keyhole Limpet Hemocyanin-using EDC as described by the manufacturer (Pierce), followed by injection of the coupled immunogen into rabbits. The antibodies were first precipitated from the serum (#6076) with an equal volume of saturated ammonium chloride followed by resuspension and dialysis against PBS. Affinity purification was carried out using a peptide column prepared by coupling the antigenic peptide to CNBr-activated Sepharose (Pharmacia) as described by the manufacturer. The antibodies were then bound to the peptide column and washed with 2 m KCl-PBS. Elution was carried out with 20 ml S m Nal (in 1 mM sodium thiosulfate), which was dialyzed immediately against PBS.

20

25.

To generate monoclonal antibodies, female Balb/c mice were immunized with 50 ug AgDH-2 or AgDH-3. Additional mice were immunized with 25 to 50 ug AgDH-2 or AgDH-3 that had been combined with an equal molar ratio of mAb 61F3B, a monoclonal antibody with specific reactivity to GST. A third group of mice were immunized with SDS polyacrylamide gel slices containing AgDH-2 or AgDH-3. The immunogen for each group of mice was prepared in complete Freund's adjuvant, with subsequent boosts (25 ug antigen in incomplete Freund's) at about 21 day intervals. Cell lines producing monoclonal antibodies were isolated as follows. Briefly a single cell suspension was formed by grinding immunized mouse spleen in serum free RPMI 1640, supplemented with 2mM L-glutamine, 1mM sodium pyruvate, 100 units/ml penicillin, and 100 μ g/ml streptomycin (RPMI) (Gibco, Canada). The cell suspension was filtered through

30

sterile 70-mesh Nitex cell strainer (Becton Dickinson, Parsippany, New Jersey), and washed twice by centrifuging at 200 g for 5 minutes and resuspending the pellet in 20 ml serum free RPMI. Thymocytes taken from three naive Balb/c mice were prepared in this manner.

5

NS-1 myeloma cells kept in log phase in RPMI with 11% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, were centrifuged at 200 g for 5 minutes, and the pellet was washed twice as described in the foregoing paragraph. After washing, each cell suspension was brought to a final volume of 10 ml in serum free RPMI, and 10 μ l was diluted 10:100. Twenty μ l of each dilution was removed, mixed with 20 μ l 0.4% trypan blue stain in 0.85% saline (Gibco), loaded onto a hemacytometer and counted.

10

15

20

Two x 10^8 spleen cells were combined with 4 x 10^7 NS-1 cells, centrifuged, and the supernatant was aspirated. The cell pellet was dislodged by tapping the tube and 2 ml of 37° C PEG 1500 (50% in 75mM Hepes, pH 8.0) (Boehringer Mannheim) was added with stirring over the course of 1 minute, followed by adding 14 ml of serum free RPMI over 7 minutes. An additional 16 ml RPMI was added and the cells were centrifuged at 200 g for 10 minutes. After discarding the supernatant, the pellet was resuspended in 200 ml RPMI containing 15% FBS, $100~\mu$ M sodium hypoxanthine, 0.4μ M aminopterin, 16μ M thymidine (HAT) (Gibco), 25 units/ml IL-6 (Mallinckrodt, Folcrost, Pennsylvania), and 1.5~x 10^{6} thymocytes/ml. The suspension was dispensed into ten 96-well flat bottom tissue culture plates at $200~\mu$ l/well. Cells in plates were fed 3 to 4 times between fusing and screening by aspirating approximately half the medium from each well with an 18 G needle and replenishing plating medium described above except containing 10 units/ml IL-6 and lacking thymocytes.

25

30

Fusions were screened when cell growth reached 60-80% confluency (day 7 to 9) by ELISA on AgDH2 versus AgDH3. Immunlon 4 plates (Dynatech, Cambridge, MA) were coated at 4°C overnight with 100 ng/well protein in 30mM carbonate buffer, pH 9.6. Plates were blocked with 100 μ g/well 0.5% fish skin gelatin in PBS for one hour at 37°C, washed 3 times with PBS, 0.05% Tween 20 (PBST) and 50 μ l culture supernatant is added. After incubation at 37°C for 30 minutes, and washing as described above, 50 μ l of horseradish peroxidase conjugated

10

15

20

25

goat anti-mouse IgG(fc) (Jackson ImmunoResearch. West Grove, PA) diluted 1:10,000 in PBST was added. Plates were incubated as above, washed 4 times with PBST and 100 μ l substrate consisting of 100 μ g/ml of tetramethylbenzidine and 0.15 μ l/ml H₂O₂ in 100mM sodium acetate, pH 5.5, was added. The color reaction was stopped in 5-10 minutes with the addition of 50 μ l of 15% H₂SO₄. A₄₉₀ was read on a plate reader.

Fifty three pools of hybridomas that were positive in an ELISA were screened for the ability to immunoblot or immunoprecipitate MCCS from a mouse testes cell lysate. Immunoblot analysis using the mouse testes extract is described in Example 6. Immunoprecipitations was performed as follows. A six percent SDS polyacrylamide gel was run and transferred to Immobilon-PVDF in 192 mM glycine, 25 mM Tris base, 0.1% SDS, 20% methanol, then blocked for 1 hour in 5% powdered nonfat milk, 20 mM Tris ph 7.5, 100 mM NaCl 0.1% Tween 20, and cut into the appropriate number of strips. The primary antibody (well supernatant) was diluted in the above block solution and incubated for one hour at room temperature, washed four times in block minus milk, incubated in goat anti-mouse IgG (H+L) HRP (BioRad #170-6516), washed again in block solution minus milk, transfered to NEN Renaissance ECL reagent and developed for 5 minutes.

Immunoprecipitation was performed as follows. Fifty μ l of hybridoma supernatant was incubated for one hour on ice with 300 μ g of testes cell lysate prepared as described in Example 6. Thirty μ l of a 50% slurry of protein A agarose (Pierce, Rockford, IL), prebound to a rabbit anti-mouse bridging antibody (5 μ g/reaction) (Pierce) was added and incubated at 4°C with rocking. The immune complexes were washed three times in lysis buffer and the antigen/antibody complex eluted by boiling in SDS sample buffer (2% SDS, 20 mM Tris pH 6.8, 20% glycerol, 0.001% bromphenol blue). The resulting supernatant was separated on a 6% SDS polyacrylamide gel and transferred to Immobilon-PVDF (Millipore) and an immunoblot was performed using affinity purified rabbit anti-Ag DH2 polyclonal antiserum. Four hybridomas were cloned and characterized in immunoblots, immunoprecipitations and in immunoprecipitation/kinase assays as described in Example 6. The four hybridoma cell lines were designated 224B, 224C (ATCC HB

30

12233), 224F (ATCC HB 12234) and 224G. All four monoclonal antibodies recognized MCCS1 by immunoblot and immunoprecipitation.

Example 6

MCCS1 associated protein kinase activity was immunoprecipitated using the MCCS1-specific polyclonal antibodies described in Example 5.

Extracts were made from fresh testes tissue isolated from Balb/c mice. Minced testes were homogenized on ice with 10-15 strokes of a tight fitting dounce homogenizer in Lysis Buffer (50 mM NaPO4, pH 7.2; 0.5% TritonX-100; 2 mM EDTA; 2 mM EGTA; 25 mM NaF; 25 mM 2-glycerophosphate; 1 mM phenylmethylsulfonyl fluoride [PMSF]; I μ g/ml leupeptin; I μ g/ml pepstatin A; 2 mM DTT) and incubated on ice for 30 minutes. The lysate was centrifuged at 13.000xg rpm for 10 minutes at 4° C in a TL-100 table-top ultracentrifuge (Beckman) to remove unbroken cells and other insoluble material. Aliquots of cell lysate were snap frozen in liquid N₂ and stored at -70°C. Five hundred ug of testes extract was incubated with either 5 ug of affinity purified anti-AgDH-2 polyclonal antibody or 5 ug purified rabbit IgG (Zymed, So. San Francisco, CA) in 1 ml of Lysis buffer for one hour on ice in microcentrifuge tubes. Thirty μl of protein A sepharose beads (Repligen, Cambridge, MA) (washed in Lysis buffer) were added to the extracts, and then incubated for an additional 30 minutes at 4° C on a rocking platform. The immune complex/Protein A sepharose beads were washed four times with 1 ml of Lysis buffer, one time with 1 ml Kinase Buffer (25 mM Hepes pH 7.7; 50 mM KCl: 10 mM MgCl₂; 0.1% NP-40; 2% glycerol; 1 mM DTT), and then incubated in 20 ul Kinase Buffer with 10 μ Ci ATP [50 Ci/mmol]) for 20 minutes at 37°C. The kinase reactions were stopped with 20 µl 2X SDS sample buffer and heated to 100° C prior to separation on 6% polyacrylamide gels. Gels were fixed in 20% methanol/7% Acetic acid, and then dried onto Whatman 3MM paper prior to autoradiography. While little or no phosphorylation was evident in control immunoprecipitations, immunoprecipitations using anti-AgDH-2 antibody contained two major phosphorylated bands at approximately 300 kD and approximately 180 kD. In addition, there were several minor phosphorylation products, including one which comigrated with the MCCS1 protein itself as demonstrated by Western blot analysis

5

10

15

20

25

(see Example 8 for Western blot description.) Phosphoaminoacid analysis of the approximately 300 kD protein identified the presence of phosphoserine residues. Addition of 5 ug of AgDH-2 (but not AgDH-3) dramatically reduced or eliminated the MCCS1-associated kinase activity found in the immunoprecipitates.

5

10

15

20

25

Example 7

The expression pattern of MCCS1 in various human tissues was examined by Northern blot hybridization.

Nylon membranes containing 2 μ g of size-fractionated polyA+ RNA from a variety of human tissue sources were obtained from Clontech Laboratories, Inc., and the hybridization protocol supplied by the manufacturer was followed precisely, except that the final wash was performed at 55° C, rather than 50° C, to minimize the possibility of cross-hybridization to related sequences. The ³²P-labelled DNA hybridization probe used was generated by PCR. A DNA encoding the COOH-terminal 30% of MCCS1 α was used as a template to amplify a 1.3 kb fragment in the presence of ³²P-dCTP using primers 279-3 5'TGGATGATGACAGCTGTGTC 3' (SEQ ID NO: 21) and 279-6 5'TGTAGTCGCTGCTCAATGTC3' (SEQ ID NO: 22).

Results of the Northern blots show that MCCS1 is expressed as an approximately 9 kb mRNA in a wide variety of human tissues. Testis tissue contains the highest level of MCCS1 mRNA, though the transcript is also expressed in small intestine, ovary, prostate, thymus, spleen, heart, peripheral blood lymphocytes, colon, brain, placenta, skeletal muscle, kidney and pancreas.

Expression of MCCS1 mRNA in human cancer cell lines was also examined using a human cancer cell line RNA blot obtained from Clonetech. The RNA blot contained RNA from the cell lines HL-60 (promyelocytic leukemia), HeLa K-562 (cervical carcinoma), (chronic myelogenous leukemia), MOLT-4 (lymphoblastic leukemia), Raji (Burkitt's lymphoma), SW480 (colorectal adenocarcinoma), A549 (lung carcinoma), and G361 (melanoma). Northern blot analysis was performed as directed by the manufacturer with hybridization being carried out at 65°C using a 2.0kb Kpnl-Sall fragment of the MCCS1 partial clone HFB2. Expression was observed in the HL-60, HeLa, K-562, Raji, SW480, and G361 cell lines with the highest level of expression occurring in the G361 cell line.

10

15

20

25

30

Detectable but low levels of expression were observed in the MOLT-4 and A549 cell lines.

Example 8

The expression of MCCS1 mRNA and protein in normal and irradiated mouse testes and in mouse embryos was examined by *in situ* hybridization, immunostaining and/or immunoblotting.

In situ Hybridization

Normal and irradiated mouse testes were harvested from male Balb/c mice. The tissues were sectioned at 6µm thickness, picked up on Superfrost Plus® (VWR Scientific) slides and allowed to air-dry at room temperature overnight. Sections were stored at -70° C if not immediately used. The tissue sections were fixed in 4% paraformaldehyde (Sigma) in PBS for 20 minutes at 4° C, dehydrated (70%, 95%, 100% ethanol) for 1 minute at 4° C in each grade, then allowed to air dry for 30 minutes at room temperature. The slides were acetylated in a solution of 0.25% (v/v) acetic anhydride (Sigma)/0.1M triethanolamine pH 8.0 for 10 minutes at room temperature with stirring, rinsed in 0.2X SSC for 10 minutes at room temperature with stirring, and dehydrated and air dried as described above. The tissues were hybridized in situ with digoxigenin-labeled single-stranded mRNA generated from murine MCCS1 DNA by in vitro RNA transcription incorporating digoxigen-UTP (Boehringer Mannheim). The labeled riboprobes (see sequence in SEQ ID NO: 27) (lµg/section) and diethylpyrocarbonate (depc)-treated water were added to hybridization buffer with a final concentration of 50% formamide, 0.3 M NaCl. 20 mM Tris pH 7.5, 10% dextran sulfate, 1X Denhardt's solution, 100 mM dithiothreitol (DTT) and 5 mM EDTA, and 20 μ l of the solution was applied to each section and covered with a sterile, RNase-free 22 x 22 cover slip. The mRNA in both the section and the probe solution was denatured by heating the slides to 85° C for 10 minutes in an oven. Hybridization was carried out overnight (12-16 hours) at 50° C.

After hybridization, sections were washed for 1 hour at room temperature in 4X SSC/10 mM DTT, then for 30 minutes at 50° C in 50% formamide/2X SSC/10 mM DTT, 30 minutes at 37° C in a solution of 500 mM

10

15

20

25

30

NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5 (NTE buffer), 30 minutes at 37° C in a bath of 10 µg/mL RNase A (Boehringer Mannheim) in NTE buffer, 15 minutes at 37° C in NTE buffer, 15 minutes at room temperature in 2X SSC, 15 minutes at room temperature in 0.1X SSC, and 2 minutes at room temperature in 100 mM Tris-HC1, 150 mM NaCl, pH 7.5 (Buffer 1). To detect the labeled riboprobes, the sections were blocked for 30 minutes at room temperature in a solution of 5% normal sheep serum (Harlan Bioproducts for Science, Indianapolis, IN) and 0.3% Triton X-100 (Sigma) in Buffer 1 with gentle stirring, after which 150 µl/section of sheep α Digoxigenin-gold conjugate (Goldmark Biologicals, Philipburg, Pa) was applied to the tissues and incubated for 2 hours at room temperature. The slides were then washed three times for 5 minutes in Buffer 1, five times for 3 minutes in sterile deionized water, the excess liquid blotted off the slide and 2 drops each of silver enhancing and initiating solution (Goldmark Biologicals) applied to each section. The chemical reaction was allowed to proceed for 23 minutes at room temperature, then the sections were rinsed thoroughly in sterile deionized water, counterstained in Nuclear Fast Red (Vector), rinsed again in sterile deionized water, air dried overnight at room temperature and mounted with Cytoseal 60 (VWR).

In both normal and irradiated mouse testes signal was observed in the cytoplasm of spermatogonia and spermatocytes. The expression level in irradiated testis was not increased over that seen in normal testis.

Immunostaining

Testis tissue from normal male Balb/c mice was sectioned at 6 μ m thickness, picked up on Superfrost Plus® (VWR Scientific) slides and allowed to airdry at room temperature overnight. Sections were stored at -70° C if not immediately used. The sections were fixed in cold (4° C) acetone for 10 minutes at room temperature; once the slides were removed from the acetone the reagent was allowed to evaporate from the sections. Each tissue section was blocked with 150 μ l of a solution of 30% normal rat serum (Harlan Bioproducts). 5% normal goat serum (Vector Laboratories) and 1% bovine serum albumin (BSA) (Sigma) in 1X TBS for 30 minutes at room temperature. After blocking, the solution was gently blotted from the sections and anti-AgDH-3 and anti-AgDH-2 polyclonal antibodies and preimmune sera from the same rabbits were diluted 1:50 and 1:100 in the blocking solution and

10

15

20

100 µl applied to each tissue section and incubated for 30 minutes at 37° C. The antibody solution was blotted gently from the sections and unbound antibody removed from the sections by washing the slides 3 times for 5 minutes each in 1X TBS. The excess TBS was blotted from the slide and 100 µl of the biotinylated goat anti-rabbit antibody contained in the Elite Rabbit IgG Vectastain ABC kit (Vector), prepared according to the product insert, were applied to each section and incubated for 15 minutes at 37° C. After incubation, the slides were washed 2 times in 1X TBS for 5 minutes in each wash. Next, 100 µl of streptavidin-gold conjugate (Goldmark Biologicals) diluted 1:100 in a solution containing 5% normal rat serum and 1% BSA was applied to each section and incubated for 1 hour at room temperature. The slides were then washed 3 times in 1X TBS for 5 minutes each wash, and 100 µl of 1% glutaraldehyde (Sigma) in TBS buffer was applied to the slides for 5 minutes at room temperature. The slides were then washed 3 times for 5 minutes each in TBS, then 4 times in sterile deionized water for 3 minutes each. The excess liquid was blotted from each slide and 2 drops each of silver enhancing and initiating solution (Goldmark Biologicals) were applied to each section. The chemical reaction was allowed to proceed for 13 minutes at room temperature, then the sections were rinsed thoroughly in sterile deionized water, counterstained in Nuclear Fast Red (Vector), rinsed again in sterile deionized water, air dried overnight at room temperature and mounted with Cytoseal 60 (VWR).

Signal was detected in the spermatogonia and primary spermatocytes with both of the polyclonal antibodies, but not with the preimmune sera from the same animals.

Immunoblotting

25

30

Freshly obtained mouse testicles were minced with razor blades in cold PBS, and a cell suspension was generated using a loose fitting dounce homogenizer. This cell suspension was then boiled with an equal volume of 2X SDS sample buffer. Fifty ug aliquots of each extract were separated on 6% polyacrylamide gels, transferred onto Immobilon membranes (Millipore, Bedford, MA) and analyzed for anti-MCCS1-reactivity using the affinity purified antibodies in Example 5, and HRP-conjugated goat anti-rabbit secondary antibody and the Renaissance Enhanced Chemiluminescence kit (Dupont/NEN, Boston, MA). Extracts prepared from fresh

10

15

20

mouse testis contain a high molecular weight species (about 294 kD) that was recognized by both affinity-purified antiserum. No reactivity against this protein was seen with either of the preimmune sera. Importantly, the signal obtained from each affinity purified sera was specifically blocked after pre-incubation of the antibody with the corresponding immunogen.

In summary, high levels of MCCS1 mRNA and protein are detected in mouse testis in the spermatogonia and primary spermatocytes, cells that are in the early stages of meiosis. This suggests that MCCS1 plays an important role in meiotic cell division. Meiosis is a specialized form of cell division that produces germ cells in higher eukaryotes. There are two major characteristics of meiosis that distinguish it from mitosis. Whereas mitotic cell division results in genetically identical cells containing two of each chromosome, meiotic cell division results in cells containing one of each chromosome. Early in meiosis, during the "reduction division" process, sister chromatids pair and undergo reciprocal recombination at some regions. During this process, these cells are exposed to DNA strand breaks. It is likely that the cellular response to the DNA strand breaks during meiosis is similar to the cellular response found in non-germ cells in response to IR-induced DNA damage. This interpretation is further substantiated by studies that demonstrate the MEC1 is upregulated 10 to 20 fold during sporulation, indicating an important role for MCCS1 during meiosis in addition to its role in DNA repair.

Example 9

In order to identify the cells within the developing mouse testis that express MCCS1. Western blot analysis of MCCS1 expression within populations of meiotic cells was performed. Extracts of purified pachytene spermatocytes, round spermatids, condensing spermatids, and epididymal sperm cells were examined for MCCS1 expression as described above in Example 8.

Pachytene spermatocytes, round, and condensing spermatids were prepared from decapsulated testes of adult mice by sequential dissociation with collagenase and trypsin-DNase 1. The cells were separated into discrete populations by sedimentation velocity at unit gravity in 2-4% BSA gradients in Enriched Krebs Ringer Bicarbonate Medium (EKRB). The pachytene spermatocyte and round

spermatid populations were each at least 85% pure, while the condensing spermatid population was about 40-50% pure (contaminated primarily with enucleated residual bodies and some round spermatids). Sperm were obtained from the cauda epididymides. Purified populations of spermatogenic cells were dissolved directly in SDS-sample buffer containing 40 mM DTT, heated to 100° C for 5 minutes, and the amount of protein in each sample determined by the Amido-Black procedure.

The highest levels of MCCS1 protein were found in pachytene spermatocytes, with the level dropping significantly in round spermatids. MCCS1 protein levels were barely detectable lower in the condensing spermatid population, and this may reflect the presence of round spermatids in the preparation (see above). No MCCS1 protein was detected in epididymal sperm. The Western analysis thus corroborates the immunocytochemical data, and suggests a role for MCCS1 in meiotic cells.

Example 10

15

20

25

30

5

10

Substrates of MCCS1 and proteins that interact with MCCS1 (for example, members of the cell cycle checkpoint pathway and proteins that localize MCCS1 in cells) may be identified by various assays.

A. Identification of Substrates

Substrates of MCCS1 may be identified by incorporating test compounds in assays for kinase activity. MCCS1 kinase is resuspended in 20 μ l kinase buffer (25mM Hepes pH7.4, 25mM KCl, 10mM MgCl2, 1mM DTT, 2% glycerol, 0.1% NP40, 0.5mM ATP, 10 uCl gamma 32 P-ATP) and incubated for 30 minutes, either in the presence or absence of 4 μ g test compound (e.g., casein, histone H1, or appropriate substrate peptide). Reactions are separated on 12% PAGE gels and dried onto Whatman paper prior to autoradiography. Moles of phosphate transferred by the kinase to the test compound are measured by autoradiography or scintillation counting. Transfer of phosphate indicates that the test compound is a substrate of the kinase.

The protein PHAS-1 has been identified as an *in vitro* substrate of ATR (Example 4). PHAS-1 is a heat and acid-stable protein that phosphorylated at several sites *in vivo* in response to insulin and growth factors. PHAS-1 binds to the mRNA

10

15

20

binding factor, EIF-4E, and prevents translation of capped mRNAs. Phosphorylation of PHAS-1 at a specific serine residue results in dissociation of PHAS-1 form EIF-4E and thus releasing the inhibition of translation of capped mRNAs. This mechanism allows for a rapid synthesis of protein in response to a particular stimulus. PHAS-1 may be phosphorylated by several protein kinases in vivo including a protein kinase that is sensitive to rapamycin. Since the rapamycinsensitive protein kinase, FRAP, is related to ATR, it would be reasonable to assume that there might be an overlap in substrate specificity between FRAP and ATR and that PHAS-1 is a substrate for both of these protein kinases in vitro. To test this hypothesis, ATR that was immunoprecipitated from a mouse testes cell extract or Histagged ATR purified from baculovirus-infected SF9 cells (Example 4) was incubated with 10 µg PHAS-I (Stratagene) in kinase buffer (25 mM Hepes pH 7.4, 25 mM KC1, 10 mM MgC1₂, 1 mM DTT, 0.1% NP-40), 10 μ M ATP and 10 μ Ci₃₂P γ ATP for 20 minutes at 37°C. Since phosphorylated PHAS-1 was known to bind to phosphocellulose paper, the reaction was spotted onto phosphocellulose spin columns and centrifuged at 2500 x g, washed twice with 0.5 ml of 75 mM phosphoric acid and once with 0.5 ml absolute ethanol. The phosphocellulose disks were then transferred to scintillation vials and the counts per minutes incorporated into the PHAS-1 proteins were recorded. ATR readily phosphorylated PHAS-1 whereas negative controls showed little or no PHAS-1 phosphorylation. To map which residue is phosphorylated, the following peptides representing PHAS-1 sequences containing serine and threonine residues were synthesized.

Peptide PH-1

MSGGSSCQTPSRAIPATRR (SEQ ID NO: 36)

Peptide PH-2

GDYSTTPGGTLFSTTPGGTRR (SEQ ID NO: 37)

Peptide PH-3

ECRNSPVTKTRR (SEQ ID NO: 38)

Peptide PH-4

30 GVTSPSSDEPRR (SEQ ID NO: 39)

Peptide PH-5

MEASQSHLRR (SEQ ID NO: 40)

15

20

25

30

4

Peptide PH-6

RRNSPEDKRAGG (SEQ ID NO: 41)

Peptide PH-7

GEESQFEMDIRR (SEQ ID NO: 42)

These peptides are tested in the same kinase reaction to determine which peptide(s) is (are) phosphorylated by ATR. The peptide(s) are then used as substrate for ATR or MCCS1 in assays such as described in Example 11 to identify modulators.

The same kinase reaction was also used to determine if proteins such as histone H1 (Upstate Biotechnology, Inc., Waltham, NY) and myelin basic protein (Gibco BRL, Gaithersburg, MD) which are known to be substrates of other protein kinases are substrates of MCCS1 and ATR. No phosphorylation of histone H1 or myelin basic protein was observed under the conditions of the assay. Moreover, a peptide from p53 known to be a substrate of DNA-PK was also not phosphorylated in the assay.

B. Identification of Interacting Proteins

Interacting proteins may be identified by the following assays.

A first assay contemplated by the invention is a two-hybrid screen. The two-hybrid system was developed in yeast [Chien et al., Proc. Natl. Acad. Sci. USA, 88: 9578-9582 (1991)] and is based on functional in vivo reconstitution of a transcription factor which activates a reporter gene. Specifically, a polynucleotide encoding a protein that interacts with MCCS1 is isolated by: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding a first fusion of part or all of MCCS1 and either the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a library of second hybrid DNA sequences encoding second fusions of part or all of putative MCCS1 binding proteins and the DNA binding domain or activating domain of the transcription factor which is not incorporated in the first fusion; detecting binding of an MCCS1 interacting protein to MCCS1 in a particular host cell by detecting the production of reporter gene product in the host cell; and isolating second hybrid DNA sequences encoding the interacting protein from the particular

host cell. Presently preferred for use in the assay are a *lexA* promoter to drive expression of the reporter gene, the *lacZ* reporter gene, a transcription factor comprising the *lexA* DNA binding domain and the GAL4 transactivation domain, and yeast host cells.

5

Other assays for identifying proteins that interact with MCCS1 may involve immobilizing MCCS1 or a test protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the amount of label bound. Bound label indicates that the test protein interacts with MCCS1.

10

15

Another type of assay for identifying MCCS1 interacting proteins involves immobilizing MCCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling a test protein with a compound capable of exciting the fluorescent agent, contacting the immobilized MCCS1 with the labelled test protein, detecting light emission by the fluorescent agent, and identifying interacting proteins as test proteins which result in the emission of light by the florescent agent. Alternatively, the putative interacting protein may be immobilized and MCCS1 may be labelled in the assay.

Example 11

20

Modulators of MCCS1 include MCCS1 variants and other molecules. The modulators may affect MCCS1 kinase activity, its localization in the cell, and/or its interaction with members of the cell cycle checkpoint pathway. Presently preferred regions of MCCS1 which are targets for mutation or the development of selective modulators include the following four regions: the MCCS1 α amino terminal effector domain (amino acids 1 to 1081 of SEQ ID NO: 31), the MCCS1 β amino terminal effector domain (amino acids 1 to 1150 of SEQ ID NO: 33), the MCCS1 α rad3+ domain (amino acids 1082 to 2082 of SEQ ID NO: 31), the MCCS1 β rad3+ domain (amino acids 1151 to 2151 of SEQ ID NO: 33), the MCCS1 α PIK domain (amino acids 2083 to 2410 of SEQ ID NO: 31), and the MCCS1 β PIK domain (amino acids 2152 to 2480 of SEQ ID NO: 33).

25

30

MCCS1 variants having mutations in the kinase domain may be useful as a radiosensitizing agents. Mutations specifically contemplated by the invention are.

replacement of the MCCS1 α aspartic acid at amino acid 2241, the asparagine at 2246, and the aspartic acid at 2260 of SEQ ID NO: 31 with alanine or methionine, and the corresponding mutations in MCCS1 β . Analogous mutations in the rad3+ gene resulted in yeast hypersensitive to radiation. In addition, mutations in the kinase domain of ATM are found in patients with AT, a disease that causes radiation sensitivity.

Furthermore, combinatorial libraries, peptide and peptide mimetics, defined chemical entities, oligonucleotides, and natural product libraries may be screened for activity as modulators in assays such as those described below.

15

20

10

5

For example, an assay for identifying modulators of MCCS1 kinase activity involves incubating an MCCS1 kinase preparation in kinase buffer with gamma-³²P-ATP and an exogenous kinase substrate, both in the presence and absence of a test compound, and measuring the moles of phosphate transferred to the substrate. For example, 2 µl of the 50 mM imidazole elution pool is added to kinase buffer. (See Example 6.) The reactions are incubated at 37°C for 20 min and samples are analyzed by SDS-PAGE prior to autoradiography or Western analysis. An increase in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the test compound is an activator of said MCCS1 kinase. Conversely, a decrease in the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in presence of the test compound compared to the moles of phosphate transferred to the substrate in the absence of the test compound indicates that the modulator is an inhibitor of said MCCS1 kinase.

25

Moreover, assays for identifying compounds that modulate interaction of MCCS1 with other proteins may involve: transforming or transfecting appropriate host cells with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA-binding domain and an activating domain; expressing in the host cells a first hybrid DNA sequence encoding a first fusion of part or all of MCCS1 and the DNA binding domain or the activating domain of the transcription factor; expressing in the host cells a second hybrid DNA sequence encoding part or all of a protein that interacts with MCCS1 and the DNA binding domain or activating domain of the transcription factor which is not

10

15

20

25

30

incorporated in the first fusion; evaluating the effect of a test compound on the interaction between MCCS1 and the interacting protein by detecting binding of the interacting protein to MCCS1 in a particular host cell by measuring the production of reporter gene product in the host cell in the presence or absence of the test compound; and identifying modulating compounds as those test compounds altering production of the reported gene product in comparison to production of the reporter gene product in the absence of the modulating compound. Presently preferred for use in the assay are a *lexA* promoter to drive expression of the reporter gene, the *lacZ* reporter gene, a transcription factor comprising the *lexA* DNA binding domain and the GALA transactivation domain, and yeast host cells.

Another type of assay for identifying compounds that modulate the interaction between MCCS1 and an interacting protein involves immobilizing MCCS1 or a natural MCCS1 interacting protein, detectably labelling the nonimmobilized binding partner, incubating the binding partners together and determining the effect of a test compound on the amount of label bound wherein a reduction in the label bound in the present of the test compound compared to the amount of label bound in the absence of the test compound indicates that the test agent is an inhibitor of MCCS1 interaction with protein. Conversely, an increase in the bound in the presence of the test compound compared to the amount label bound in the absence of the compound indicates that the putative modulator is an activator of MCCS1 interaction with the protein.

Yet another method contemplated by the invention for identifying compounds that modulate the binding between MCCS1 and an interacting protein involves immobilizing MCCS1 or a fragment thereof on a solid support coated (or impregnated with) a fluorescent agent, labelling the interacting protein with a compound capable of exciting the fluorescent agent, contacting the immobilized MCCS1 with the labelled interacting protein in the presence and absence of a test compound, detecting light emission by the fluorescent agent, and identifying modulating compounds as those test compounds that affect the emission of light by the fluorescent agent in comparison to the emission of light by the fluorescent agent in the absence of the test compound. Alternatively, the MCCS1 interacting protein may be immobilized and MCCS1 may be labelled in the assay.

Example 12

Cell based complementation assays for identifying modulators of MCCS1 or ATM are described below.

In one type of assay, host cells (for example, esr1-1 yeast cells) are transformed with MCCS1-encoding DNA as is described in Example 4. The esr1-1 yeast strain is normally sensitive to treatment with ultraviolet (UV) light, but esr1-1 yeast cells expressing MCCS1 or ATR are no longer sensitive to treatment with UV light. The transformed yeast cells are exposed to test compounds and the effect of the test compounds on UV sensitivity of the transformed host cell is determined. Test compounds that are inhibitors of MCCS1 or ATR activity restore UV sensitivity to the MCCS1 transformed esr1-1 cells. Alternatively, esr1-1 tel1 double mutant yeast cells are used as host cells instead of esr1-1 yeast cells. The TEL1 gene is homologous to ATM and the TEL1 mutation is described in Morrow, et al., Cell, 82:831-840 (1995). The invention also specifically contemplates that the esr1-1 or esr1-1 tel1 double mutant yeast host cells may be transformed with ATM-encoding DNA (SEQ ID NO: 34).

In an alternative embodiment, the assays include clastogenic agents or events instead of treatment with UV light (e.g., IR, hydroxyurea, or DNA damaging agents). Appropriate host cells for use in such embodiments would be those that are sensitive to the alternative clastogenic agents or events.

Another type of complementation assay involves the use of mammalian host cells such as cell lines derived from cells of AT patients. As described above for yeast cells, the mammalian cells are transfected with DNA encoding MCCS1, ATR, or ATM and then exposed to test compounds. Test compounds that are inhibitors of MCCS1, ATR, or ATM activity will restore the phenotype of the untransformed host cell (e.g., sensitivity to IR).

The above assays can be used to identify compounds that inhibit activity of MCCS1, ATR, and ATM or compounds that inhibit activity of only one of the enzymes.

In an alternative type of assay, the yeast or mammalian host cells are transformed with DNA encoding chimeric polypeptides including various combinations of MCCS1 and ATM domains. MCCS1 and ATM show structural

15

10

5

20

10

15

20

25

30

similarities, and chimeric polypeptides which comprise portions of MCCS1 and ATM are useful in elucidating active sites and binding domains of both MCCS1 and ATM. Polynucleotides encoding the chimeras can be prepared by standard molecular biology techniques known to the skilled worker and as exemplified herein. The chimeric polypeptides are expressed in host cells and modulators of the chimeras can be identified by the assays disclosed herein.

Example 13

MCCS1 and ATM are both involved in meiosis I checkpoints. Since MCCS1 is demonstrated herein to have kinase activity, assays were performed to determine if ATM possessed kinase activity. To determine the kinase activity of ATM, ATM was immunoprecipitated from MRC-5 fibroblasts (ATCC #171-CCL) with polyclonal antisera, 6076. MRC-5 cells are human lung embryonal diploid fibroblasts. MRC-5 cells were obtained from the ATCC at passage 19 and maintained in Minimal Essential Medium supplemented with 10% fetal bovine serum, 100 units/ml penicillin, 100 mg/ml streptomycin, and 100 mM MEM non-essential amino acids. Media and media supplements were obtained through Gibco Life Technologies. Cell lines were maintained in a water-saturated 37°C incubator with 5% C.

MRC-5 cell extracts were prepared by lysis of a 10cm plate of log-phase cells in 0.5 ml of Lysis Buffer I (50 mM NaPO₄, pH 7.2; 0.5% TritonX-100; 2 mM EDTA; 2 mM EGTA; 25 mM NaF; 25 mM 2-glycerophosphate; 1 mM phenylmethylsulfonyl fluoride [PMSF]; 1 μ g/ml leupeptin 1 μ g/ml pepstatin A; 2 mM DTT) on ice. Cells were scraped from plates using a rubber spatula then sonicated in a cup horn sonicator (Sonifier 250, Branson Ultrasonics Corp., Danbury, CT) at 100% output for 90 seconds. Lysates were then clarified in a 4°C microfuge for 2 minutes. Preclearing was done by adding 10 μ g purified rabbit IgG (Zymed) and 30 μ l Protein A Agarose slurry (Pierce) followed by incubation at 4°C for 60 minutes while rocking. To the precleared lysates, 10 μ g of affinity purified 6076 antisera (or 10 μ g 6076 pre-blocked with 0.04 mg P45 peptide for 30 min.) was added and incubated on ice for 60 minutes. Immunoprecipitates were collected by addition of

30 μ l Protein A agarose slurry and incubated with rocking at 4°C for 30 minutes followed by four washes in Lysis Buffer I.

Kinase reactions were carried out by washing the immunoprecipitations once with kinase buffer (25 μ M Hepes pH 7.7; 50 mM KCl; 10 mM MgCl₂; 0.1% NP-40; 2% glycerol; 1 mM DTT), followed by incubation in 20 μ l of Kinase Buffer containing 10 μ M ATP + 10 μ Ci γ ³²P-ATP [50 Ci/mmol] for 20 minutes at 37°C. Reactions were stopped by the addition of 20 μ l 2X SDS sample buffer and boiled for 5 minutes prior to separation on 6% SDS polyacrylamide gels. The gels were dried and exposed to x-ray film (Kodak, XAR-5) at -80°C overnight.

10

5

10 cm plates of log-phase MRC-5 cells were washed once with PBS then incubated in Dulbeco's Modified Eagle Medium (minus methionine) containing 2% dialyzed fetal bovine serum for 30 minutes. Cells were labeled by adding 200 μ Ci³⁵S-methionine (1000 Ci/mmol TRAN³⁵S-LABEL, ICN Radiochemicals) for 2 hours. Labeled cells were then washed once with PBS and frozen at -80°C prior to immunoprecipitation.

15

The incubation of the immunoprecipitated complexes in kinase buffer produced a phosphorylated product with a molecular weight of approximately 350,000 that co-migrated with ATM in polyacrylamide gels.

20

Similar results were obtained for ATR immune complexes immunoprecipitated with anti-AgDH-2 (MCCS1) polyclonal antisera of Example 5. ATR and ATM thus appear to be able to self-phosphorylate or associate with a protein kinase.

25

To determine the role of ATR and ATM in meiosis, immunostaining techniques on surface spreads of mouse spermatocytes were utilized to localize ATR and ATM to meiotic chromosomes. Antibodies recognizing ATR and ATM were utilized with mouse antibodies against Corl. Corl is a component of axial/lateral elements of synapsing chromosomes [Dobson et al., J. Cell Sci., 107:2749-2760 (1994)]. Corl chromosomal staining appears when the axial elements begin to form between the sister chromatids of each homolog in leptonema of meiotic prophase, prior to the initiation of synapsis. As homologous bivalents synapse, the axial elements from the two homologs align and a central element forms between them, completing the structure called the synaptonemal complex (SC).

10

15

20

25

30

When short stretches of Corl begin to appear prior to any evidence of synapsis, neither ATR nor ATM is detectable. As homologs start synapsis, both proteins were seen at pairing forks; however, the location and behavior of the two proteins differed markedly. In normal zygotene nuclei, the stage during which homologs synapse, ATR was present in small amounts and transiently at discrete foci along the asynapsed (unpaired) axes. As homologs synapse, ATR disappeared from these locations. However, at regions delayed in synapsis, often seen near the proximal ends of autosomal bivalents, there was an accumulation of ATR foci along the unsynapsed axes. ATR was detected at similar locations on the two axial elements. In nuclei where an entire autosome fails to find its homologous pairing partner, ATR foci were detected along the entire lengths of these asynapsed axis. In males, where the X chromosome has no homolog, ATR foci were localized along the unpaired axis.

ATM was also visualized as foci and was first detected during zygonema as homologs synapse, but ATM localization was different than ATR. ATM was first observed along synapsed axes when homologous autosomal axial elements come into contact. However, during mid-pachynema, after autosomal synapsis has been completed, ATM foci appeared on the X chromosome axis. ATM localization persisted on fully synapsed bivalents into pachynema, a substage that lasts 3 days in mouse oocytes and 6 days in mouse spermatocytes. During pachynema, the number of foci drops gradually, stabilizing briefly in mid-pachynema before eventually disappearing mid-to late pachynema. Thus, ATR and ATM protein kinases play important and complementary roles at distinct stages in meiosis I.

The involvement of ATR appears to be transient during early meiotic prophase while the role of ATM appears to be more prolonged. However, both ATR and ATM coordinate the various events of meiotic prophase by performing similar checkpoint functions.

The foregoing illustrative examples relate to presently preferred embodiments of the invention and numerous modifications and variations thereof are expected to occur to those skilled in the art. Thus only such limitations as appear in the appended claims should be placed upon the scope of the present invention.

PCT/US96/19337

-51-

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
 - (i) APPLICANT: ICOS Corporation
 - (ii) TITLE OF INVENTION: Cell Cycle Checkpoint PIK-Related Kinase Materials and Methods
 - (iii) NUMBER OF SEQUENCES: 42
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 - (B) STREET: 6300 Sears Tower, 233 S. Wacker Dr.

 - (C) CITY: Chicago
 (D) STATE: Illinois
 - (E) COUNTRY: USA (F) ZIP: 60606

 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Noland, Greta E.
 - (B) REGISTRATION NUMBER: 35,302
 - (C) REFERENCE/DOCKET NUMBER: 27866/33607
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (312) 474-6300 (B) TELEFAX: (312) 474-0448 (C) TELEX: 25-3856
- (2) INFORMATION FOR SEQ ID NO:1:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7621 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (vii) IMMEDIATE SOURCE:
 - (B) CLONE: pBSHFB2HT2-27
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 333..7559
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTTGI	rgaa	GA (GAAT	GTTT	TA C	ACTC	TTGT	T AG	TGAA	GTTT	TTA	CTTT	AAA	AGTO	AATCGT	•	60
CAAGG	FATT	TA (GCAA	ATGA	AT T	AGCA	CTTC	g Ga	ATAT	CTTG	TTT	'ATTT	TAA	ATCT	TTTTTG	•	120
TTTAT	TTC	AA A	AGAA'	TTCA	GT A	ATTG	GATC	А ТА	ACGA	GACT	TCT	GCGG	ATT	GCAG	CAACTC		180
CCTCC	TGT	CA :	TTTG'	TTAC.	AC A	AGAA	AATC	T GT	GAAG	TCAT	CTG	TTCA	TTA	TATT	TTCTTT		240
TTAAA	LAGC:	AA (GAGT	CCTG	CT A	TTTT	TGGG	G TA	CTCA	CAAA	AGA	ATTA	TTA	CAAC	TTTTTG		300
AAGAC	TTG	GT 1	TTAC	CTCC	A TA	GAAG.	AAAT(G TG	ATG Met 1	GGT Gly	CAT His	GCT Ala	GTG Val 5	Glu	TGG Trp		353
CCA G Pro V	TG (GTC Val 10	ATG Met	AGC Ser	CGA Arg	TTT	TTA Leu 15	AGT Ser	CAA Gln	TTA Leu	GAT Asp	GAA Glu 20	CAC His	ATG Met	GGA Gly		401
TAT T Tyr L	TA (eu (25	CAA Gln	TCA Ser	GCT Ala	CCT Pro	TTG Leu 30	CAG Gln	TTG Leu	ATG Met	AGT Ser	ATG Met 35	Gln	AAA Lys	TTA Leu	GAA Glu		449
TTT A Phe I 40	TT (GAA Glu	GTC Val	ACT Thr	TTA Leu 45	TTA Leu	ACG Thr	GTT Val	CTT Leu	ACT Thr 50	CGT Arg	ATT Ile	ATT	GCA Ala	ATT Ile 55		497
GTG T	TT 1 he F	rrr Phe	AGA Arg	AGG Arg 60	CAA Gln	GAA Glu	CTC Leu	TTA Leu	CTT Leu 65	TGG Trp	CAG Gln	ATA Ile	GGT Gly	TGT Cys 70	GTT Val		54 5
CTG C	TA C	GAG Glu	TAT Tyr 75	GGT Gly	AGT Ser	CCA Pro	AAA Lys	ATT Ile 80	AAA Lys	TCC Ser	CTA Leu	GCA Ala	ATT Ile 85	AGC Ser	TTT Phe		593
TTA A	CA G	SAA Slu 90	CTT Leu	TTT Phe	CAG Gln	CTT Leu	GGA Gly 95	GGA Gly	CTA Leu	CCA Pro	GCA Ala	CAA Gln 100	CCA Pro	GCT Ala	AGC Ser		641
ACT T	TT T he F 05	TTC Phe	AGC Ser	TCA Ser	TTT Phe	TTG Leu 110	GAA Glu	TTA Leu	TTA Leu	AAA Lys	CAC His 115	CTT Leu	GTA Val	GAA Glu	ATG Met	•	689
GAT AG Asp Tl 120	CT G	ASP	CAA Gln	TTG Leu	AAA Lys 125	CTC Leu	TAT Tyr	GAA Glu	GAG Glu	CCA Pro 130	TTA Leu	TCA Ser	AAG Lys	CTG Leu	ATA Ile 135	•	737
AAG AG Lys Tì	CA C hr L	TA Seu	TTT Phe	CCC Pro 140	TTT Phe	GAA Glu	GCA Ala	GAA Glu	GCT Ala 145	TAT Tyr	AGA Arg	AAT Asn	ATT Ile	GAA Glu 150	CCT Pro	•	785
GTC TA	AT T yr L	eu	AAT Asn 155	ATG Met	CTG Leu	CTG Leu	GAA Glu	AAA Lys 160	CTC Leu	TGT Cys	GTC Val	ATG Met	TTT Phe 165	GAA Glu	GAC Asp	8	333
GGT GT Gly Va	al L	TC eu 70	ATG Met	CGG Arg	CTT Leu	AAG Lys	TCT Ser 175	GAT Asp	TTG Leu	CTA Leu	AAA Lys	GCA Ala 180	GCT Ala	TTG Leu	TGC Cys	8	381
CAT THE	TA C eu L 85	TG eu	CAG Gln	TAT Tyr	TTC Phe	CTT Leu 190	AAA Lys	TTT Phe	GTG Val	CCA Pro	GCT Ala 195	GGG Gly	TAT Tyr	GAA Glu	TCT Ser	9	929
GCT TT Ala Le 200	ΓA C eu G	AA ln	GTC Val	AGG Arg	AAG Lys 205	GTC Val	TAT Tyr	GTG Val	AGA Arg	AAT Asn 210	ATT Ile	TGT Cys	AAA Lys	GCT Ala	CTT Leu 215	9	7 7

				GGA Gly 220												1025
				TTG Leu												1073
				CAA Gln												1121
				CGT Arg					_	_						1169
				ACT Thr												1217
				AGT Ser 300												1265
				AGT Ser												1313
				GTC Val												1361
				ATG Met												1409
TCC Ser 360	AAG Lys	AAG Lys	AAA Lys	CCT Pro	TCT Ser 365	GTA Val	GTG Val	ATA Ile	ACT Thr	TGG Trp 370	ATG Met	TCA Ser	TTG Leu	GAT Asp	TTT Phe 375	1457
				CTT Leu 380												1505
AAA Lys	CGG Arg	ACT Thr	GGA Gly 395	GGC Gly	AAC Asn	ATT Ile	GAT Asp	AAG Lys 400	GTG Val	GTG Val	AAA Lys	ATT Ile	TAT Tyr 405	GAT Asp	GCT Ala	1553
TTG Leu	ATT Ile	TAT Tyr 410	ATG Met	CAA Gln	GTA Val	AAC Asn	AGT Ser 415	TCA Ser	TTT Phe	GAA Glu	GAT Asp	CAT His 420	ATC Ile	CTG Leu	GAA Glu	1601
GAT Asp	TTA Leu 425	TGT Cys	GGA Gly	ATG Met	CTC Leu	TCA Ser 430	CTT Leu	CCA Pro	TGG Trp	ATT Ile	TAT Tyr 435	TCC Ser	CAT His	TCT Ser	GAT Asp	1649
				AAG Lys												1697
AGC Ser	TGT Cys	AGG Arg	ATT Ile	TCA Ser 460	GAT Asp	AGC Ser	TAT Tyr	TCA Ser	CCA Pro 465	CAG Gln	GCA Ala	CAA Gln	TCA Ser	CGA Arg 470	TGT Cys	1745
GTG Val	TTT Phe	CTT Leu	CTG Leu 475	ACT Thr	CTG Leu	TTT Phe	Pro	AGA Arg 480	AGA Arg	ATA Ile	TTC Phe	CTT Leu	GAG Glu 485	TGG Trp	AGA Arg	1793

						GCC Ala							-	_		1841
						TTT Phe 510										1889
						ATT Ile										1937
						TTT Phe										1985
						TAT Tyr										2033
						GAC Asp										2081
						TCT Ser 590										2129
						AAA Lys										2177
						CAT His										2225
						AAA Lys										2273
						GAT Asp										2321
						TTG Leu 670										2369
CTT Leu 680	TTT Phe	GTC Val	TTA Leu	AGA Arg	ATG Met 685	AAG Lys	GAA Glu	GCA Ala	TAT Tyr	ACA Thr 690	CAT His	GCC Ala	CAA Gln	ATA Ile	TCA Ser 695	2417
						GAT Asp										2465
						GAT Asp										2513
						AAG Lys										2561
						GTT Val 750										2609

						AAA Lys										269	57
						ACA Thr										270	05
						GAT Asp										275	53
						GCC Ala										280	01
						TTA Leu 830										284	19
						TCT Ser										289	∍7
						GAG Glu										294	15
						TCT Ser										299	3
						GAA Glu										304	1
						TTG Leu 910										308	39
GAA Glu 920	CAC His	TAT Tyr	CAA Gln	CAG Gln	GTT Val 925	TTT Phe	AAT Asn	GGT Gly	TTG Leu	TCA Ser 930	ATA Ile	CTT Leu	GCC Ala	TCA Ser	TTT Phe 935	313	37
						TAT Tyr										318	35
GAA Glu	CTG Leu	ATG Met	GCT Ala 955	GAT Asp	TAT Tyr	TTA Leu	CAA Gln	CCC Pro 960	AAA Lys	TTG Leu	TTG Leu	GGC Gly	ATT Ile 965	TTG Leu	GCT Ala	323	3
TTT Phe	TTT Phe	AAC Asn 970	ATG Met	CAG Gln	TTA Leu	CTG Leu	AGC Ser 975	TCT Ser	AGT Ser	GTT Val	GGC Gly	ATT Ile 980	GAA Glu	GAT Asp	AAG Lys	328	31
AAA Lys	ATG Met 985	GCC Ala	TTG Leu	AAC Asn	AGT Ser	TTG Leu 990	ATG Met	TCT Ser	TTG Leu	ATG Met	AAG Lys 995	TTA Leu	ATG Met	GGA Gly	CCC Pro	332	9
AAA Lys 1000	His	GTC Val	AGT Ser	TCT Ser	GTG Val 1005	AGG Arg	GTG Val	AAG Lys	ATG Met	ATG Met 1010	Thr	ACA Thr	CTG Leu	AGA Arg	ACT Thr 1015	337	7
GGC Gly	CTT Leu	CGA Arg	TTC Phe	AAG Lys 1020	qaA	GAT Asp	TTT Phe	CCT Pro	GAA Glu 1025	Leu	TGT Cys	TGC Cys	AGA Arg	GCT Ala 1030	Trp	342	5

											•		•			
GAC Asp	TGC Cys	TTT Phe	GTT Val 103	Arg	TGC Cys	CTG Leu	GAT Asp	CAT His 104	Ala	TGT Cys	CTG Leu	GGC Gly	TCC Ser 1045	Leu	CTC Leu	3473
			Ile			TTG Leu		Pro					Gln			3521
GAA Glu	ACT Thr 1065	Ala	GCT Ala	ATC Ile	TTC Phe	CAC His	Tyr	CTC Leu	ATA Ile	ATT Ile	GAA Glu 1079	Asn	AGG Arg	GAT Asp	GCT Ala	3569
GTG Val 1080	Gln	GAT Asp	TTT Phe	CTT Leu	CAT His 1085	GAA Glu	ATA Ile	TAT Tyr	TTT Phe	TTA Leu 1090	Pro	GAT Asp	CAT His	CCA Pro	GAA Glu 1095	3617
					Ala	GTT Val				Tyr					Ser	3665
GAG Glu	AGC Ser	ACT Thr	GAT Asp 1115	Leu	CAG Gln	ACA Thr	ACT Thr	CTT Leu 1120	Gln	CTC Leu	TCT Ser	ATG Met	AAG Lys 1125	Ala	ATT Ile	3713
			Asn			GTT Val		Ile					Ser			3761
GAA Glu	ACC Thr 1145	Leu	TAT Tyr	AAA Lys	AAT Asn	CAG Gln 115	Glu	AAA Lys	CTG Leu	ATA Ile	AAG Lys 1155	Tyr	GCA Ala	ACA Thr	GAC Asp	3809
AGT Ser 1160	Glu	ACA Thr	GTA Val	GAA Glu	CCT Pro 1165	ATT Ile	ATC Ile	TCA Ser	CAG Gln	TTG Leu 1170	Val	ACA Thr	GTG Val	CTT Leu	TTG Leu 1175	3857
					Ala	AAC Asn				Arg					Glu	3905
TGT Cys	TTA Leu	GGG Gly	GAA Glu 1199	Leu	GGG Gly	GCG Ala	ATA Ile	GAT Asp 1200	Pro	GGT Gly	CGA Arg	TTA Leu	GAT Asp 1205	Phe	TCA Ser	3953
			Thr			AAA Lys		Phe					Gly			4001
		Ser				GGA Gly 1230	Leu					Thr				4049
	Ala					AGC Ser					Ser					4097
					Ser	ATT Ile				Arg					Asn	4145
GGC Gly	CCA Pro	GGT Gly	CAC His 1275	Gln	TTG Leu	TGG Trp	AGG Arg	AGA Arg 1280	Phe	CCT Pro	GAG Glu	CAT His	GTT Val 1285	Arg	GAA Glu	4193
ATA Ile	CTA Leu	GAA Glu 1290	Pro	CAT His	CTA Leu	TAA naA	ACC Thr 1295	Arg	TAC Tyr	AAG Lys	AGT Ser	TCT S r 1300	Gln	AAG Lys	TCA Ser	4241

WO 97/18323

		Trp					Lys					Ser			GGT Gly	4289
	Asn					Ser					Gly				ACA Thr 1335	4337
AAG Lys	GTT Val	CGA Arg	CAT His	GAT Asp 134	Leu	GCC Ala	AGT Ser	AAA Lys	ATT Ile 134	Phe	ACC Thr	TGC Cys	TGT Cys	AGC Ser 135	Ile	4385
ATG Met	ATG Met	AAG Lys	CAT His 135	Asp	TTC Phe	AAA Lys	GTG Val	ACC Thr 136	Ile	TAT Tyr	CTT Leu	CTT Leu	CCA Pro 136	His	ATT Ile	4433
CTG Leu	GTG Val	TAT Tyr 1370	Val	TTA Leu	CTG Leu	GGT Gly	TGT Cys 1375	Asn	CAA Gln	GAA Glu	GAT Asp	CAG Gln 138	Gln	GAG Glu	GTT Val	4481
TAT Tyr	GCA Ala 138	Glu	ATT Ile	ATG Met	GCA Ala	GTT Val 139	Leu	AAG Lys	CAT His	GAC Asp	GAT Asp 139	Gln	CAT His	ACC Thr	ATA Ile	4529
AAT Asn 140	Thr	CAA Gln	GAC Asp	ATT Ile	GCA Ala 140	Ser	GAT Asp	CTG Leu	TGT Cys	CAA Gln 1410	Leu	AGT Ser	ACA Thr	CAG Gln	ACT Thr 1415	4577
					Asp		CTC Leu			Trp					Phe	4625
				Ala			TGT Cys		His					Arg		4673
			Ser				ACT Thr 1455	Val					Tyr			4721
GTA Val	ACC Thr 1465	Arg	T TT Phe	CTA Leu	GAC Asp	CTC Leu 1470	ATA Ile	CCC Pro	CAG Gln	GAT Asp	ACT Thr 1475	Leu	GCA Ala	GTA Val	GCT Ala	4769
TCC Ser 148	Phe	CGC Arg	TCC Ser	AAA Lys	GCA Ala 1489	Tyr	ACA Thr	CGA Arg	GCT Ala	GTA Val 1490	Met	CAC His	TTT Phe	GAA Glu	TCA Ser 1495	4817
TTT Phe	ATT Ile	ACA Thr	GAA Glu	AAG Lys 1500	Lys	CAA Gln	AAT Asn	ATT Ile	CAG Gln 1509	Glu	CAT His	CTT Leu	GGA Gly	TTT Phe 1510	Leu	4865
CAG Gln	AAA Lys	TTG Leu	TAT Tyr 1515	Ala	GCT Ala	ATG Met	CAT His	GAA Glu 1520	Pro	GAT Asp	GGA Gly	GTG Val	TCC Ser 1525	Gly	GTC Val	4913
AGT Ser	GCA Ala	ATT Ile 1530	Arg	AAG Lys	GCA Ala	GAA Glu	CCA Pro 1535	Ser	CTA Leu	AAA Lys	GAA Glu	CAG Gln 1540	Ile	CTT Leu	GAA Glu	4961
CAT His	GAA Glu 1545	Ser	CTT Leu	GGC Gly	TTG Leu	CTG Leu 1550	AGG Arg	GAT Asp	GCC Ala	ACT Thr	GCT Ala 1555	Cys	TAT T yr	GAC Asp	AGG Arg	5009
GCT Ala 1560	Ile	CAG Gln	CTA Leu	GAA Glu	CCA Pro 1565	Asp	CAG Gln	ATC Ile	ATT Ile	CAT His 1570	Tyr	CAT His	GGT Gly	GTA Val	GTA Val 1575	5057

					CTT Leu)					Thr					Val	5105
				Ala	AAC Asn				Trp					Asn		5153
			Glu		GCT Ala			Leu					Leu			5201
		Leu			GAT Asp		Lys					Ser				5249
	Gln				TCA Ser 1645	Ala					Ile					5297
GAC Asp	TCA Ser	CTG Leu	AAA Lys	CTA Leu 1660	GTG Val	AGA Arg	GCA Ala	GAA Glu	CAA Gln 1665	Ile	GTA Val	CCT Pro	CTT Leu	TCA Ser 1670	Ala	5345
				Arg	GGC Gly				Arg					Ile		5393
			Met		TGT Cys			Glu					Pro			5441
		Ser			GAC Asp		Ser					Leu				5489
	Arg				ACC Thr 1725	Gln					Ala					5 5 37
					GCT Ala					Asn					Tyr	5585
				Gly	GAA Glu				Gln					Ala		5633
			His		CAG Gln			Tyr					Asn			5681
		Arg			GAA Glu		Tyr					Lys				5729
	Lys				CAC His 1805	Gln					Leu					5 <i>777</i>
					GAA Glu					Pro					Met	5825
				Arg	GCT Ala				Val					Glu		5873

	GCT Ala		Phe					Ile					Lys			592
	GCG Ala 1869	Cys					Glu					Tyr				5969
	TAT Tyr 0					Pro					Asn					6017
	GGT Gly				Arg					His					Leu	6065
	TAT Tyr			Gln					Ser					Leu		6113
	TGG Trp		Asp					Ala					Lys			6161
	TCC Ser 1945	Asp					Arg					Lys				6209
	ATC Ile O					Asn					Tyr					6257
	TTT Phe				Ile					His					Val	6305
	GTT Val			Asp					Gln					Tyr		6353
	CAA Gln		Met					Ala					Ser			6401
	CGT Arg 2025	Val					Glu					Ala				6449
	AAA Lys					Phe					Thr					6497
AAG Lys	CTT Leu	CTA Leu	GAA Glu	TTG Leu 2060	Cys	AAT Asn	AAA Lys	CCG Pro	GTG Val 2065	Glu	ATT Ile	CTT Leu	GCT Ala	TCT Ser 2070	Leu	6545
CAG Gln	AAA Lys	CCA Pro	AAG Lys 2075	Lys	ATT Ile	TCT Ser	TTA Leu	AAA Lys 2080	Gly	TCA Ser	GAT Asp	GGA Gly	AAG Lys 2085	Phe	TAC Tyr	6593
ATC Ile	ATG Met	ATG Met 2090	Cys	AAG Lys	CCA Pro	AAA Lys	GAT Asp 2095	Asp	CTG Leu	AGA Arg	AAG Lys	GAT Asp 2100	Cys	AGA Arg	CTA Leu	6641
ATG Met	GAA Glu 2105	Phe	AAT Asn	TCC Ser	TTG Leu	ATT Ile 2110	Asn	AAG Lys	TGC Cys	TTA Leu	AGA Arg 2115	Lys	GAT Asp	GCA Ala	GAG Glu	6689

Ser Arg Arg Arg Glu Leu 2120 2125	His Ile Arg Th	A TAT GCA GTT ATT Tyr Ala Val Ile 2130	CCA CTA 673 Pro Leu 2135
AAT GAT GAA TGT GGG ATT Asn Asp Glu Cys Gly Ile 2140	ATT GAA TGG GTG Ile Glu Trp Val 214	. Asn Asn Thr Ala	GGT TTG 678 Gly Leu 2150
AGA CCT ATT CTG ACC AAA Arg Pro Ile Leu Thr Lys 2155	CTA TAT AAA GAA Leu Tyr Lys Glu 2160	AAG GGA GTG TAT Lys Gly Val Tyr 216	Met Thr
GGA AAA GAA CTT CGC CAG Gly Lys Glu Leu Arg Gln 2170	TGT ATG CTA CCA Cys Met Leu Pro 2175	AAG TCA GCA GCT Lys Ser Ala Ala 2180	TTA TCT 6883 Leu Ser
GAA AAA CTC AAA GTA TTC Glu Lys Leu Lys Val Phe 2185	CGA GAA TTT CTC Arg Glu Phe Leu 2190	CTG CCC AGG CAT Leu Pro Arg His 2195	CCT CCT 6929 Pro Pro
ATT TTT CAT GAG TGG TTT Ile Phe His Glu Trp Phe 2200 2205	Leu Arg Thr Phe	CCT GAT CCT ACA Pro Asp Pro Thr 2210	TCA TGG 6977 Ser Trp 2215
TAC AGT AGT AGA TCA GCT Tyr Ser Ser Arg Ser Ala 2220	TAC TGC CGT TCC Tyr Cys Arg Ser 222	Thr Ala Val Met	TCA ATG 7025 Ser Met 2230
GTT GGT TAT ATT CTG GGG Val Gly Tyr Ile Leu Gly 2235	CTT GGA GAC CGT Leu Gly Asp Arg 2240	CAT GGT GAA AAT His Gly Glu Asn 224	Ile Leu
TTT GAT TCT TTG ACT GGT Phe Asp Ser Leu Thr Gly 2250	GAA TGC GTA CAT Glu Cys Val His 2255	GTA GAT TTC AAT Val Asp Phe Asn 2260	TGT CTT 7121 Cys Leu
TTC AAT AAG GGA GAA ACC Phe Asn Lys Gly Glu Thr 2265	TTT GAA GTT CCA Phe Glu Val Pro 2270	GAA ATT GTG CCA Glu Ile Val Pro 2275	TTT CGC 7169 Phe Arg
CTG ACT CAT AAT ATG GTT Leu Thr His Asn Met Val 2280 2285	AAT GGA ATG GGT Asn Gly Met Gly	CCT ATG GGA ACA Pro Met Gly Thr 2290	GAG GGT 7217 Glu Gly 2295
CTT TTT CGA AGA GCA TGT Leu Phe Arg Arg Ala Cys 2300	GAA GTT ACA ATG Glu Val Thr Met 230	Arg Leu Met Arg	GAT CAG 7265 Asp Gln 2310
CGA GAG CCT TTA ATG AGT Arg Glu Pro Leu Met Ser 2315	GTC TTA AAG ACT Val Leu Lys Thr 2320	TTT CTA CAT GAT Phe Leu His Asp 2325	Pro Leu
GTG GAA TGG AGT AAA CCA Val Glu Trp Ser Lys Pro 2330	GTG AAA GGG CAT Val Lys Gly His 2335	TCC AAA GCG CCA Ser Lys Ala Pro 2340	CTG AAT 7361 Leu Asn
GAA ACT GGA GAA GTT GTC . Glu Thr Gly Glu Val Val . 2345	AAT GAA AAG GCC Asn Glu Lys Ala 2350	AAG ACC CAT GTT Lys Thr His Val 2355	CTT GAC 7409 Leu Asp
ATT GAG CAG CGA CTA CAA Ile Glu Gln Arg Leu Gln 2360 2365	GGT GTA ATC AAG Gly Val Ile Lys	ACT CGA AAT AGA Thr Arg Asn Arg 2370	GTG ACA 7457 Val Thr 2375
GGA CTG CCG TTA TCT ATT (Gly Leu Pro Leu Ser Ile (2380	GAA GGA CAT GTG Glu Gly His Val 238	His Tyr Leu Ile	CAA GAA 7505 Gln Glu 2390

-61-

GCT ACT GAT GAA AAC TTA CTA TGC CAG ATG TAT CTT GGT TGG ACT CCA
Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro
2395

TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA AGTAAAAAAA
7609

TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA AGTAAAAAAA 76
Tyr Met

AAAAAAAA AA 7621

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2409 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser
1 10 15

Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu 20 25 30

Met Ser Met Gln Lys Leu Glu Phe Ile Glu Val Thr Leu Leu Thr Val

Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu 50 55 60

Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile 65 70 75 80

Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly 85 90 95

Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu 100 105 110

Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu 115 120 125

Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu 130 135 140

Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys 145 150 155 160

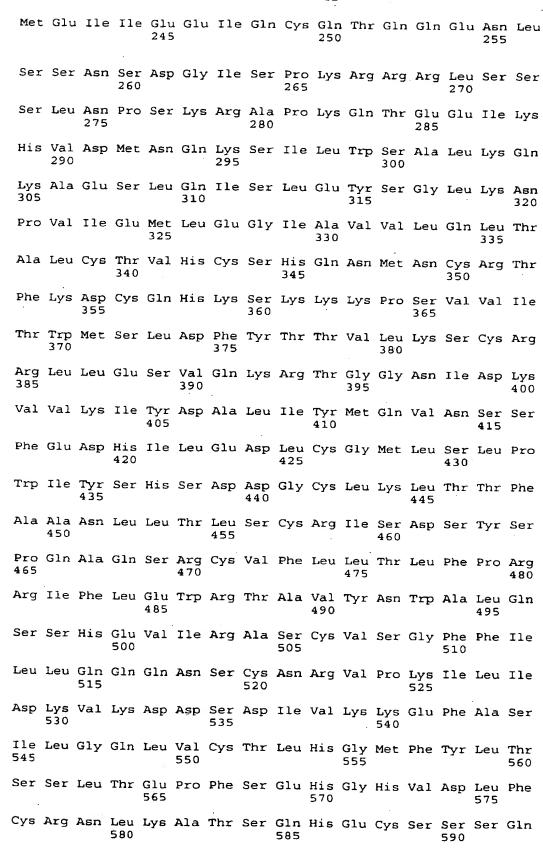
Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp 165 170 175

Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe 180 185 190

Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val 195 200 205

Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp 210 215 220

Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser 225 230 235 240



Leu	Lys	Ala 595	Ser	Val	Cys	Lys	Pro 600	Phe	Leu	Phe	Leu	Leu 605	Lys	Lys	Ly
Ile	Pro 610	Ser	Pro	Val	Lys	Leu 615	Ala	Phe	Ile	Asp	Asn 620	Leu	His	His	Let
Cys 625	Lys	His	Leu	Asp	Phe 630	Arg	Glu	Asp	Glu	Thr 635	Asp	Val	Lys	Ala	Va.1 640
Leu	Gly	Thr	Leu	Leu 645	Asn	Leu	Met	Glu	Asp 650	Pro	Asp	Lys	Asp	Val 655	Arg
Val	Ala	Phe	Ser 660	Gly	Asn	Ile	Lys	His 665	Ile	Leu	Glu	Ser	Leu 670	Asp	Ser
Glu	Asp	Gly 675	Phe	Ile	Lys	Glu	Leu 680	Phe	Val	Leu	Arg	Met 685	Lys	Glu	Ala
Tyr	Thr 690	His	Ala	Gln	Ile	Ser 695	Arg	Asn	Asn	Glu	Leu 700	Lys	Asp	Thr	Let
Ile 705	Leu	Thr	Thr	Gly	Asp 710	Ile	Gly	Arg	Ala	Ala 715	Lys	Gly	Asp	Leu	Va] 720
Pro	Phe	Ala	Leu	Leu 725	His	Leu	Leu	His	Cys 730	Leu	Leu	Ser	Lys	Ser 735	Ala
Ser	Val	Ser	Gly 740	Ala	Ala	Tyr	Thr	Glu 745	Ile	Arg	Ala	Leu	Val 750	Ala	Ala
Lys	Ser	Val 755	Lys	Leu	Gln	Ser	Phe 760	Phe	Ser	Gln	Tyr	Lys 765	Lys	Pro	Ile
Cys	Gln 770	Phe	Leu	Val	Glu	Ser 775	Leu	His	Ser	Ser	Gln 780	Met	Thr	Ala	Leu
Pro 785	Asn	Thr	Pro	Cys	Gln 790	Asn	Ala	Asp	Val	Arg 795	Lys	Gln	Asp	Val	Ala 800
His	Gln	Arg	Glu	Met 805	Ala	Leu	Asn	Thr	Leu 810	Ser	Glu	Ile	Ala	Asn 815	Val
Phe	Asp	Phe	Pro 820	Asp	Leu	Asn	Arg	Phe 825	Leu	Thr	Arg	Thr	Leu 830	Gln	Val
Leu	Leu	Pro 835	Asp	Leu	Ala	Ala	Lys 840	Ala	Ser	Pro	Ala	Ala 845	Ser	Ala	Leu
Ile	Arg 850	Thr	Leu	Gly	Lys	Gln 855	Leu	Asn	Val	Asn	Arg 860	Arg	Glu	Ile	Let
11e 865	Asn	Asn	Phe	Lys	Tyr 870	Ile	Phe	Ser	His	Leu 875	Val	Cys	Ser	Cys	Ser 880
Lys	Asp	Glu	Leu	Glu 885	_	Ala	Leu	His	Tyr 890	Leu	Lys	Asn	Glu	Thr 895	Glu
Ile	Glu	Leu	Gly 900	Ser	Leu	Leu	Arg	Gln 905	Asp	Phe	Gln	Gly	Leu 910	His	Asn
Glu	Leu	Leu 915	Leu	Arg	Ile	Gly	Glu 920	His	Tyr	Gln	Gln	Val 925	Phe	Asn	Gly
Leu	Ser 930	Ile	Leu	Ala	Ser	Phe 935	Ala	Ser	Ser	Asp	Asp 940	Pro	Tyr	Gln	Gly

- Pro Arg Asp Ile Ile Ser Pro Glu Leu Met Ala Asp Tyr Leu Gln Pro 945 950 955 960
- Lys Leu Leu Gly Ile Leu Ala Phe Phe Asn Met Gln Leu Leu Ser Ser 965 970 975
- Ser Val Gly Ile Glu Asp Lys Lys Met Ala Leu Asn Ser Leu Met Ser 980 985 990
- Leu Met Lys Leu Met Gly Pro Lys His Val Ser Ser Val Arg Val Lys 995 1000 1005
- Met Met Thr Thr Leu Arg Thr Gly Leu Arg Phe Lys Asp Asp Phe Pro 1010 1015 1020
- Glu Leu Cys Cys Arg Ala Trp Asp Cys Phe Val Arg Cys Leu Asp His 1025 1030 1035 1040
- Ala Cys Leu Gly Ser Leu Leu Ser His Val Ile Val Ala Leu Leu Pro 1045 1050 1055
- Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr Leu 1060 1065 1070
- Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile Tyr 1075 1080 1085
- Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu Gln 1090 1095 1100
- Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr Leu 1105 1110 1115 1120
- Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg Ile 1125 1130 1135
- His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu Lys 1140 1145 1150
- Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile Ser 1155 1160 1165
- Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser Gln 1170 1175 1180
- Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile Asp 1185 1190 1195 1200
- Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp Phe 1205 1210 1215
- Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu Leu 1220 1225 1230
- Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala 1235 1240 1245
- Pro Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp 1250 1255 1260
- Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg Arg 1265 1270 1275 1280
- Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr Arg 1285 1290 1295

- Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys Pro 1300 1305 1310
- Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser 1315 1320 1325
- Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser Lys 1330 1340
- Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val Thr 1345 1350 1355 1360
- Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys Asn 1365 1370 1375
- Gln Glu Asp Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu Lys 1380 1385 1390
- His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu 1395 1400 1405
- Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu Thr 1410 1415 1420
- Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro 1425 1430 1435 1440
- His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr Val 1445 1450 1455
- Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile Pro 1460 1465 1470
- Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg 1475 1480 1485
- Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile 1490 1495 1500
- Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His Glu 1505 1510 1515 1520
- Pro Asp Gly Val Ser Gly Val Ser Ala Ile Arg Lys Ala Glu Pro Ser 1525 1530 1535
- Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg Asp 1540 1545 1550
- Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile 1555 1560 1565
- Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln Leu 1570 1580
- Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser Glu 1585 1590 1595 1600
- Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu 1605 1610 1615
- Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser 1620 1625 1630
- Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys Lys 1635 1640 1645

- Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu 1650 1660
- Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln 1665 1670 1680
- Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu Glu 1685 1690 1695
- His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln 1700 1705 1710
- Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser 1715 1720 1725
- Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser 1730 1740
- Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu 1745 1750 1755 1760
- Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr 1765 1770 1775
- Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val 1780 1785 1790
- Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala Leu 1795 1800 1805
- Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu Thr 1810 1815 1820
- Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu Leu 1825 1830 1835 1840
- Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile 1845 1850 1855
- Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu Asp 1860 1865 1870
- Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met Val 1875 1880 1885
- Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val 1890 1895 1900
- Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln 1905 1910 1915 1920
- Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala 1925 1930 1935
- Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg Asn 1940 1945 1950
- Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr Leu 1955 1960 1965
- Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile 1970 1975 1980
- Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn Ser 1985 1990 1995 2000

- Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr Ala 2005 2010 2015
- Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu Ile 2020 2025 2030
- Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val Gly 2035 2040 2045
- Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro 2050 2055 2060
- Val Glu Ile Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys 2065 2070 2075 2080
- Gly Ser Asp Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp Asp 2085 2090 2095
- Leu Arg Lys Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn Lys 2100 2105 2110
- Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg Glu Leu His Ile Arg 2115 2120 2125
- Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp 2130 2135 2140
- Val Asn Asn Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys 2145 2150 2155 2160
- Glu Lys Gly Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met Leu 2165 2170 2175
- Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe 2180 2185 2190
- Leu Leu Pro Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr 2195 2200 2205
- Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg 2210 2215 2220
- Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp 2225 2230 2235 2240
- Arg His Gly Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val 2245 2250 2255
- His Val Asp Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val 2260 2265 2270
- Pro Glu Ile Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly Met 2275 2280 2285
- Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr 2290 2295 2300
- Met Arg Leu Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys 2305 2310 2315 2320
- Thr Phe Leu His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly 2325 2330 2335
- His Ser Lys Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys 2340 2345 2350

-68-

Ala Lys Thr His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile 2355 2360

Lys Thr Arg Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His

Val His Tyr Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln 2390 2395

Met Tyr Leu Gly Trp Thr Pro Tyr Met 2405

- (2) INFORMATION FOR SEQ ID NO:3:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2835 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (vii) IMMEDIATE SOURCE: (B) CLONE: 517
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..2610
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTG Val 1	GAA Glu	GCA Ala	GCT Ala	TGG Trp 5	AAA Lys	TTG Leu	TCA	CAG Gln	TGG Trp 10	GAT Asp	TTG Leu	GTG Val	GAA Glu	AAC Asn 15	TAT Tyr	48
TTG Leu	GCA Ala	GCA Ala	GAT Asp 20	GGA Gly	AAA Lys	TCT Ser	ACA Thr	ACA Thr 25	TGG Trp	AGT Ser	GTC Val	AGA Arg	CTG Leu 30	GGA Gly	CAG Gln	96
CTA Leu	TTA Leu	TTA Leu 35	TCA Ser	GCC Ala	AAA Lys	AAA Lys	AGA Arg 40	GAT Asp	ATC Ile	ACA Thr	GCT Ala	TTT Phe 45	TAT Tyr	GAC Asp	TCA Ser	144
CTG Leu	AAA Lys 50	CTA Leu	GTG Val	AGA Arg	GCA Ala	GAA Glu 55	CAA Gln	ATT Ile	GTA Val	CCT Pro	CTT Leu 60	TCA Ser	GCT Ala	GCA Ala	AGC Ser	192
TTT Phe 65	GAA Glu	AGA Arg	GGC Gly	TCC Ser	TAC Tyr 70	CAA Gln	CGA Arg	GGA Gly	TAT Tyr	GAA Glu 75	TAT Tyr	ATT Ile	GTG Val	AGA Arg	TTG Leu 80	240
CAC His	ATG Met	TTA Leu	TGT Cys	GAG Glu 85	TTG Leu	GAG Glu	CAT His	AGC Ser	ATC Ile 90	AAA Lys	CCA Pro	CTT Leu	TTC Phe	CAG Gln 95	CAT His	288
TCT Ser	CCA Pro	GGT Gly	GAC Asp 100	AGT Ser	TCT Ser	CAA Gln	GAA Glu	GAT Asp 105	TCT Ser	CTA Leu	AAC Asn	TGG Trp	GTA Val 110	GCT Ala	CGA Arg	336
CTA Leu	GAA Glu	ATG Met 115	ACC Thr	CAG Gln	AAT Asn	TCC Ser	TAC Tyr 120	AGA Arg	GCC Ala	AAG Lys	GAG Glu	CCT Pro 125	ATC Ile	CTG Leu	GCT Ala	384

-69-

CTC Leu	CGG Arg 130	AGG Arg	GCT Ala	TTA Leu	CTA Leu	AGC Ser 135	CTC Leu	AAC Asn	AAA Lys	AGA Arg	CCA Pro 140	GAT Asp	TAC Tyr	AAT Asn	GAA Glu	432
						CTG Leu										480
						TAC Tyr										528
						GTG Val										576
						CTA Leu										624
						ACC Thr 215										672
						CTA Leu										720
						ATT Ile										768
						GAT Asp										816
						GTC Val										864
						GTT Val 295										912
						CAG Gln										960
						TCA Ser										1008
						AAT Asn										1056
						TTA Leu										1104
						ATT Ile 375										1152

			ATA Ile 390							1200
			ACA Thr							1248
			GAA Glu							1296
			GTT Val							1344
			AAA Lys							1392
			AAA Lys 470							1440
			ATT Ile							1488
			ACC Thr							1536
			TAT Tyr							1584
			AAA Lys							1632
			ATG Met 550							1680
			GAA Glu						•	1728
			CGT Arg							1776
			GAT Asp							1824
			CCT Pro							1872
			AAA Lys 630							1920
_			AAA Lys							1968

			CCT Pro 660													2016
			TGG Trp													2064
			ATG Met													2112
			CTC Leu													2160
			CTT Leu													2208
			CGC Arg 740													2256
			GGT Gly													2304
			CAG Gln													2352
			CTT Leu													2400
			AAT Asn													2448
			GAC Asp 820													2496
			ACA Thr													. 2544
			GAA Glu													2592
			CCA Pro			TGAA	ATGA	AA T	TATG	TAAA	LA GA	TATAL	GTTA			2640
ATA	TCTA	AA A	GTA	TGCA	T TI	GGTA	TGAA	тст	GTGG	TTG	TATO	TGTT	CA A	TTCT	AAAGT	2700
ACAA	CATA	L AA	TTAC	GTTC	T CA	GCAA	CTGT	TAT	TTCT	CTC	TGAT	CATT	'AA T	ТАТА	TGTAA	2760
ATAA	TATA	'AC A	TTCA	GTTA	T TA	AGAA	ATAA	ACT	GCTT	TCT	TAAT	AAAa'	A AA	AAAA	AAAA	2820
AAAA	AAAA	AA A	AAAA													2835

-72-

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 870 amino acids
 - (B) TYPE: amino acid (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4: Val Glu Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu 65 70 75 80 His Met Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val Glu Leu

Cys Phe Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met Leu Ile

His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala

Asn Phe Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val Thr Ala

Cys Leu Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr 265

Asp Lys Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys Gln Gly 280

215

Asp	Leu 290	Ile	Arg	Tyr	Ile	Val 295	Leu	His	Phe	Gly	Arg 300	Ser	Leu	Gln	Туз
Gly 305	Asn	Gln	Phe	Ile	Tyr 310	Gln	Ser	Met	Pro	Arg 315	Met	Leu	Thr	Leu	Trp 320
Leu	Asp	Tyr	Gly	Thr 325	Lys	Ser	Tyr	Glu	Trp 330	Glu	Lys	Ala	Gly	Arg 335	Sei
Asp	Arg	Val	Gln 340	Met	Arg	Asn	Asp	Leu 345	Gly	Lys	Ile	Asn	Lys 350	Val	Ile
Thr	Glu	His 355	Thr	Asn	Tyr	Leu	Ala 360	Pro	Tyr	Gln	Phe	Leu 365	Thr	Ala	Phe
Ser	Gln 370	Leu	Ile	Ser	Arg	Ile 375	Cys	His	Ser	His	Asp 380	Glu	Val	Phe	Va]
Val 385	Leu	Met	Glu	Ile	Ile 390	Ala	Lys	Val	Phe	Leu 395	Ala	Tyr	Pro	Gln	Glr 400
Ala	Met	Trp	Met	Met 405	Thr	Ala	Val	Ser	Lys 410	Ser	Ser	Tyr	Pro	Met 415	Arg
Val	Asn	Arg	Cys 420	Lys	Glu	Ile	Leu	Asn 425	Lys	Ala	Ile	His	Met 430	Lys	Lys
Ser	Leu	Glu 435	Lys	Phe	Val	Gly	Asp 440	Ala	Thr	Arg	Leu	Thr 445	Asp	Lys	Let
Leu	Glu 450	Leu	Cys	Asn	Lys	Pro 455	Val	Asp	Gly	Ser	Ser 460	Ser	Thr	Leu	Ser
Met 465	Ser	Thr	His	Phe	Lys 470	Met	Leu	Lys	Lys	Leu 475	Val	Glu	Glu	Ala	Thr 480
Phe	Ser	Glu	Ile	Leu 485	Ile	Pro	Leu	Gln	Ser 490	Val	Met	Ile	Pro	Thr 495	Leu
Pro	Ser	Ile	Leu 500	Gly	Thr	His	Ala	As n 505	His	Ala	Ser	His	Glu 510	Pro	Phe
Pro	Gly	His 515	Trp	Ala	Tyr	Ile	Ala 520	Gly	Phe	Asp	Asp	Met 525	Val	Glu	Ile
Leu	Ala 530	Ser	Leu	Gln	Lys	Pro 535	Lys	Lys	Ile	Ser	Leu 540	Lys	Gly	Ser	Asp
Gly 545	Lys	Phe	Tyr	Ile	Met 550	Met	Cys	Lys	Pro	Lys 555	Asp	Asp	Leu	Arg	Lys 560
Asp	Cys	Arg	Leu	Met 565	Glu	Phe	Asn	Ser	Leu 570	Ile	Asn	Lys	Cys	Leu 575	Arg
Lys	Asp	Ala	Glu 580	Ser	Arg	Arg	Arg	Glu 585	Leu	His	Ile	Arg	Thr 590	Tyr	Ala
Val	Ile	Pro 595	Leu	Asn	qaA	Glu	Cys 600	Gly	Ile	Ile	Glu	Trp 605	Val	Asn	Asn
Thr	Ala 610	Gly	Leu	Arg	Pro	Ile 615	Leu	Thr	Lys	Leu	Tyr 620	Lys	Glu	Lys	Gly
Val 625	Tyr	Met	Thr	Gly	Lys 630	Glu	Leu	Arg	Gln	Cys 635	Met	Leu	Pro	Lys	Ser 640

-74-

Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro

Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp

Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala

Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly

Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp

Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile 725

Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly Met Gly Pro Met

Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val Thr Met Arg Leu

Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu Lys Thr Phe Leu

His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys 795

Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr 815

His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg 820 825

Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr

Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu 855

Gly Trp Thr Pro Tyr Met 865

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA
- (vii) IMMEDIATE SOURCE:

(B) CLONE: Primer oDH15a

- (ix) FEATURE:
- (A) NAME/KEY: modified base (B) LOCATION: group(15, 18, 24, 30) (D) OTHER INFORMATION: The nucleotides at these positions are

WO 97/18323 PCT/US96/19337

-75-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5: GCAGACGGAT CCGGNWCNGA YGGNAAYHTN TAY 33 (2) INFORMATION FOR SEQ ID NO:6: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 27 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (ix) FEATURE: (A) NAME/KEY: modified_base (B) LOCATION: group (15, 18, 24)
(D) OTHER INFORMATION: The nucleotides at these positions are inosines. (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6: GCAGACGGAT CCGGNWCNGA YGGNAAY 27 (2) INFORMATION FOR SEQ ID NO:7: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 30 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (vii) IMMEDIATE SOURCE: (B) CLONE: Primer oDH16 (ix) FEATURE: (A) NAME/KEY: modified_base (B) LOCATION: 24
(D) OTHER INFORMATION: The nucleotides at these positions are inosines. (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7: GCAGACGAAT TCRCARTYRA ARTCNACRTG 30 (2) INFORMATION FOR SEQ ID NO:8: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 41 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA

(vii) IMMEDIATE SOURCE:

(B) CLONE: Primer oDH17a

-76-

- (ix) FEATURE:
 - (A) NAME/KEY: modified_base
- (B) LOCATION: group(21, 24, 27, 30)(D) OTHER INFORMATION: The nucleotides at these positions are inosines
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCAGACGGAT CCAARTTYCC NCCNRTNYTN TAYSARTGGT T

41

- (2) INFORMATION FOR SEQ ID NO:9:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 41 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA
 - (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oDH17b
 - (ix) FEATURE:
 - (A) NAME/KEY: modified_base
- (B) LOCATION: group(24, 27, 30, 33)
 (D) OTHER INFORMATION: The nucleotides at these positions are inosines.
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCAGACGAAT CCAACCAYTS RTANARNAYN GGNGGRAAYT T

41

- (2) INFORMATION FOR SEQ ID NO:10:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA
 - (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oDH18a
 - (ix) FEATURE:
 - (A) NAME/KEY: modified base
- (B) LOCATION: group (15, 18, 21, 24, 30)
 (D) OTHER INFORMATION: The nucleotides at these positions are inosines.
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCAGACGGAT CCYTNGGNYT NGGNGAYCGN CA

32

(2) INFORMATION FOR SEQ ID NO:11:

-77-

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 32 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA	
(vii) IMMEDIATE SOURCE: (B) CLONE: Primer oDH18b	
<pre>(ix) FEATURE: (A) NAME/KEY: modified_base (B) LOCATION: group(15, 18, 21) (D) OTHER INFORMATION: The nucleotides at these positions are</pre>	
inosines.	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:	
GCAGACGGAT CCYTNGGNYT NGGNGAYAGR CA	32
(2) INFORMATION FOR SEQ ID NO:12:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(ii) MOLECULE TYPE: DNA	
(vii) IMMEDIATE SOURCE: (B) CLONE: Primer oDH23	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
GACGCAGAAT TCACCAGTCA AAGAATCAAA GAG	33
(2) INFORMATION FOR SEQ ID NO:13:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(ii) MOLECULE TYPE: DNA	
(vii) IMMEDIATE SOURCE: (B) CLONE: Primer mo3	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:	
CTACAGAGCC AAGGAG	16

-78-

(2) IN	FORMATION FOR SEQ ID NO:14:	
(:	i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(i:	i) MOLECULE TYPE: DNA	
(vii	i) IMMEDIATE SOURCE: (B) CLONE: Primer mo6	
(xi	i) SEQUENCE DESCRIPTION: SEQ ID NO:14:	
TCGAGCT	TATG CTACTAGTGG GC	22
(2) INF	FORMATION FOR SEQ ID NO:15:	
(i	(A) LENGTH: 17 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA	
(vii) IMMEDIATE SOURCE: (B) CLONE: Primer oHT9-1	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:	
CCAGTAA	ACT TGCTTTC	. 17
(2) INF	ORMATION FOR SEQ ID NO:16:	• • • • • • • • • • • • • • • • • • • •
(i)) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii)) MOLECULE TYPE: DNA	
(vii)) IMMEDIATE SOURCE: (B) CLONE: Primer oHT9-4	
	SEQUENCE DESCRIPTION: SEQ ID NO:16:	
TTTGCGGC	CCC TTCCAATATC	20
(2) INFO	DRMATION FOR SEQ ID NO:17:	
(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 7440 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	

(ii) MOLECULE TYPE: cDNA

-79-

(vii) IMMEDIATE SOURCE:
 (B) CLONE: MCCSlbeta

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..7437

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATG Met 1	GGT Gly	CAT His	GCT Ala	GTG Val 5	GAA Glu	TGG Trp	CCA Pro	GTG Val	GTC Val 10	ATG Met	AGC Ser	CGA Arg	TTT Phe	TTA Leu 15	AGT Ser	48
	TTA Leu															96
ATG Met	AGT Ser	ATG Met 35	CAA Gln	AAA Lys	TTA Leu	GAA Glu	TTT Phe 40	ATT Ile	GAA Glu	GTC Val	ACT Thr	TTA Leu 45	TTA Leu	ACG Thr	GTT Val	144
CTT Leu	ACT Thr 50	CGT Arg	ATT Ile	ATT Ile	GCA Ala	ATT Ile 55	GTG Val	TTT Phe	TTT Phe	AGA Arg	AGG Arg 60	CAA Gln	GAA Glu	CTC Leu	TTA Leu	192
	TGG Trp															240
	TCC Ser															288
	CCA Pro															336
TTA Leu	AAA Lys	CAC His 115	CTT Leu	GTA Val	GAA Glu	ATG Met	GAT Asp 120	ACT Thr	GAC Asp	CAA Gln	TTG Leu	AAA Lys 125	CTC Leu	TAT Tyr	GAA Glu	384
GAG Glu	CCA Pro 130	TTA Leu	TCA Ser	AAG Lys	CTG Leu	ATA Ile 135	AAG Lys	ACA Thr	CTA Leu	TTT Phe	CCC Pro 140	TTT Phe	GAA Glu	GCA Ala	GAA Glu	432
GCT Ala 145	TAT Tyr	AGA Arg	AAT Asn	ATT Ile	GAA Glu 150	CCT Pro	GTC Val	TAT Tyr	TTA Leu	AAT Asn 155	ATG Met	CTG Leu	CTG Leu	GAA Glu	AAA Lys 160	480
CTC Leu	TGT Cys	GTC Val	ATG Met	TTT Phe 165	GAA Glu	GAC Asp	GGT Gly	GTG Val	CTC Leu 170	ATG Met	CGG Arg	CTT Leu	AAG Lys	TCT Ser 175	GAT Asp	528
TTG Leu	CTA Leu	AAA Lys	GCA Ala 180	GCT Ala	TTG Leu	TGC Cys	CAT His	TTA Leu 185	CTG Leu	CAG Gln	TAT Tyr	TTC Phe	CTT Leu 190	AAA Lys	TTT Phe	576
	CCA Pro															624
AGA Arg	AAT Asn 210	ATT Ile	TGT Cys	AAA Lys	GCT Ala	CTT Leu 215	TTG Leu	GAT Asp	GTG Val	CTT Leu	GGA Gly 220	ATT Ile	GAG Glu	GTA Val	GAT Asp	672

				CTT Leu					720
				CAA Gln					768
				TCA Ser					816
				GCA Ala 280					864
				AGC Ser					912
				TCC Ser					960
		_	_	GGA Gly				 	1008
				TCT Ser					1056
				TCC Ser 360					1104
				TAC Tyr					1152
				AAA Lys					1200
				TTG Leu					1248
				GAT Asp					1296
				GAT Asp 440					1344
				AGC Ser					1392
				GTG Val				AGA Arg 480	1440
				ACA Thr					1488

				Val											ATC Ile		1536
TTA Leu	TTG	CAG Gln 515	CAG Gln	CAG Gln	AAT Asn	TCT Ser	TGT Cys 520	AAC Asn	AGA Arg	GTT Val	CCC Pro	AAG Lys 525	ATT	CTT Leu	ATA Ile		1584
GAT Asp	AAA Lys 530	GTC Val	AAA Lys	GAT Asp	GAT Asp	TCT Ser 535	GAC Asp	ATT Ile	GTC Val	AAG Lys	AAA Lys 540	GAA Glu	TTT Phe	GCT Ala	TCT Ser		1632
ATA Ile 545	CTT Leu	GGT Gly	CAA Gln	CTT Leu	GTC Val 550	TGT Cys	ACT Thr	CTT Leu	CAC His	GGC Gly 555	ATG Met	TTT Phe	TAT Tyr	CTG Leu	ACA Thr 560		1680
AGT Ser	TCT Ser	TTA Leu	ACA Thr	GAA Glu 565	CCT Pro	TTC Phe	TCT Ser	GAA Glu	CAC His 570	GGA Gly	CAT His	GTG Val	GAC Asp	CTC Leu 575	TTC Phe		1728
TGT Cys	AGG Arg	AAC Asn	TTG Leu 580	AAA Lys	GCC Ala	ACT Thr	TCT Ser	CAA Gln 585	CAT His	GAA Glu	TGT Cys	TCA Ser	TCT Ser 590	TCT Ser	CAA Gln		1776
CTA Leu	AAA Lys	GCT Ala 595	TCT Ser	GTC Val	TGC Cys	AAG Lys	CCA Pro 600	TTC Phe	CTT Leu	TTC Phe	CTA Leu	CTG Leu 605	AAA Lys	AAA Lys	AAA Lys		1824.
ATA Ile	CCT Pro 610	AGT	CCA Pro	GTA Val	AAA Lys	CTT Leu 615	GCT Ala	TTC Phe	ATA Ile	GAT Asp	AAT Asn 620	CTA Leu	CAT His	CAT His	C TT Leu	•	1872.
TGT Cys 625	AAG Lys	CAT His	CTT Leu	GAT Asp	TTT Phe 630	AGA Arg	GAA Glu	GAT Asp	GAA Glu	ACA Thr 635	GAT Asp	GTA Val	AAA Lys	GCA Ala	GTT Val 640		1920
CTT Leu	GGA Gly	ACT Thr	TTA Leu	TTA Leu 645	AAT Asn	TTA Leu	ATG Met	GAA Glu	GAT Asp 650	CCA Pro	GAC Asp	AAA Lys	GAT Asp	GTT Val 655	AGA Arg		1968
GTG Val	GCT Ala	TTT Phe	AGT Ser 660	GGA Gly	AAT Asn	ATC Ile	AAG Lys	CAC His 665	ATA Ile	TTG Leu	GAA Glu	TCC Ser	TTG Leu 670	GAC Asp	TCT Ser		2016
GAA Glu	GAT Asp	GGA Gly 675	TTT Phe	ATA Ile	AAG Lys	GAG Glu	CTT Leu 680	TTT Phe	GTC Val	TTA Leu	AGA Arg	ATG Met 685	AAG Lys	GAA Glu	GCA Ala		2064
TAT Tyr	ACA Thr 690	CAT His	GCC Ala	CAA Gln	ATA Ile	TCA Ser 695	AGA Arg	AAT Asn	AAT Asn	GAG Glu	CTG Leu 700	AAG Lys	GAT Asp	ACC Thr	TTG Leu		2112
ATT Ile 705	CTT Leu	ACA Thr	ACA Thr	GGG Gly	GAT Asp 710	ATT Ile	GGA Gly	AGG Arg	GCC Ala	GCA Ala 715	AAA Lys	GGA Gly	GAT Asp	TTG Leu	GTA Val 720		2160
CCA Pro	TTT Phe	GCA Ala	CTC Leu	TTA Leu 725	CAC His	TTA Leu	TTG Leu	CAT His	TGT Cys 730	TTG Leu	TTA Leu	TCC	AAG Lys	TCA Ser 735	GCA Ala		2208
TCT Ser	GTC Val	TCT Ser	GGA Gly 740	GÇA Ala	GCA Ala	TAC Tyr	ACA Thr	GAA Glu 745	ATT Ile	AGA Arg	GCT Ala	CTG Leu	GTT Val 750	GCA Ala	GCT Ala		2256
AAA Lys	AGT Ser	GTT Val 755	AAA Lys	CTG Leu	CAA Gln	AGT Ser	TTT Phe 760	TTC Phe	AGC Ser	CAG Gln	TAT Tyr	AAG Lys 765	AAA Lys	CCC Pro	ATC Ile		2304

					GAA Glu									2352
					CAG Gln 790									2400
					GCT Ala				Ser					2448
					CTT Leu									2496
					GCT Ala									2544
					AAA Lys									2592
_					TAT Tyr 870									2640
					CGT Arg							Glu		2688
					CTG Leu									2736
					ATT Ile									2784
					TCA Ser									2832
			_		TCA Ser 950									2880
					TTG Leu									2928
					GAT Asp									2976
				_	GGA Gly			His				Arg		3024
		Thr			AGA Arg		Gly				Asp			3072
	Leu				GCT Ala 1030	Trp				Arg			CAT His 1040	3120

					Leu					Ile					CCT Pro 5	3168
CTT Leu	ATA Ile	CAC His	ATC Ile 106	Gln	CCT Pro	AAA Lys	GAA Glu	ACT Thr 106	Ala	GCT Ala	ATC Ile	TTC Phe	CAC His 107		CTC Leu	3216
			Asn					Gln					Glu	ATA Ile		3264
TTT Phe	TTA Leu 109	Pro	GAT Asp	CAT	CCA Pro	GAA Glu 109	Leu	AAA Lys	AAG Lys	ATA	AAA Lys 110	Ala	GTT Val	CTC Leu	CAG Gln	3312
GAA Glu 110	Tyr	AGA Arg	AAG Lys	GAG Glu	ACC Thr 111	Ser	GAG Glu	AGC Ser	ACT Thr	GAT Asp 111	Leu	CAG Gln	ACA Thr	ACT Thr	CTT Leu 1120	3360
Gln	Leu	Ser	Met	Lys 112	Ala 5	Ile	Gln	His	Glu 113	Asn O	Val	Asp	Val	CGT Arg 113	Ile 5	3408
CAT His	GCT Ala	CTT Leu	ACA Thr 1140	Ser	TTG Leu	AAG Lys	GAA Glu	ACC Thr 114	Leu	TAT Tyr	AAA Lys	AAT Asn	CAG Gln 115	GAA Glu O	AAA Lys	3456
CTG Leu	ATA Ile	AAG Lys 115	Tyr	GCA Ala	ACA Thr	GAC Asp	AGT Ser 1160	Glu	ACA Thr	GTA Val	GAA Glu	CCT Pro 116	Ile	ATC Ile	TCA Ser	3504
CAG Gln	TTG Leu 117	Val	ACA Thr	GTG Val	CTT Leu	TTG Leu 1179	Lys	GGT Gly	TGC Cys	CAA Gln	GAT Asp 1186	Ala	AAC Asn	TCT Ser	CAA Gln	3552
GCT Ala 1189	Arg	TTG Leu	CTC Leu	TGT Cys	GGG Gly 1190	Glu	TGT Cys	TTA Leu	GGG Gly	GAA Glu 1195	Leu	GGG Gly	GCG Ala	ATA Ile	GAT Asp 1200	3600
CCA Pro	GGT Gly	CGA Arg	TTA Leu	GAT Asp 1205	Phe	TCA Ser	ACA Thr	ACT Thr	GAA Glu 1210	Thr	CAA Gln	GGA Gly	AAA Lys	GAT Asp 1215	Phe	3648
ACA Thr	TTT Phe	GTG Val	ACT Thr 1220	Gly	GTA Val	GAA Glu	GAT Asp	TCA Ser 1225	Ser	TTT Phe	GCC Ala	TAT Tyr	GGA Gly 1230	TTA Leu)	TTG Leu	3696
ATG Met	GAG Glu	CTA Leu 1235	Thr	AGA Arg	GCT Ala	TAC Tyr	CTT Leu 1240	Ala	TAT Tyr	GCT Ala	GAT Asp	AAT Asn 1245	Ser	CGA Arg	GCT Ala	3744
CCA Pro	GAT Asp 1250	Ser	GCT Ala	GCC Ala	TAT Tyr	GCC Ala 1255	Ile	CAG Gln	GAG Glu	TTG Leu	CTT Leu 1260	Ser	ATT Ile	TAT Tyr	GAC Asp	3792
TGT Cys 1265	Arg	GAG Glu	ATG Met	GAG Glu	ACC Thr 1270	Asn	GGC Gly	CCA Pro	GGT Gly	CAC His 1275	Gln	TTG Leu	TGG Trp	AGG Arg	AGA Arg 1280	3840
TTT Phe	CCT Pro	GAG Glu	CAT His	GTT Val 1285	Arg	GAA Glu	ATA Ile	CTA Leu	GAA Glu 1290	Pro	CAT His	CTA Leu	AAT Asn	ACC Thr 1295	Arg	3888
TAC Tyr	AAG Lys	AGT Ser	TCT Ser 1300	Gln	AAG Lys	TCA Ser	ACC Thr	GAT Asp 1305	\mathtt{Trp}	TCT Ser	GGA Gly	GTA Val	AAG Lys 1310	AAG Lys	CCA Pro	3936

ATT Ile	TAC Tyr	TTA Leu 131	Ser	AAA Lys	TTG Leu	GGT Gly	AGT Ser 132	Asn	TTT Phe	GCA Ala	GAA Glu	TGG Trp 132	Ser	GCA Ala	TCT Ser	3984
TGG Trp	GCA Ala 133	Gly	TAT Tyr	CTT	ATT	ACA Thr 133	Lys	GTT Val	CGA Arg	CAT His	GAT Asp 134	Leu	GCC Ala	AGT Ser	AAA Lys	4032
ATT Ile 134	TTC Phe	ACC Thr	TGC Cys	TGT Cys	AGC Ser 135	Ile	ATG Met	ATG Met	AAG Lys	CAT His 1355	Asp	TTC Phe	AAA Lys	GTG Val	ACC Thr 1360	4080
ATC Ile	TAT Tyr	CTT Leu	CTT Leu	CCA Pro 1369	His	ATT Ile	CTG Leu	GTG Val	TAT Tyr 1370	Val	TTA Leu	CTG Leu	GGT Gly	TGT Cys 137	Asn	4128
CAA Gln	GAA Glu	GAT Asp	CAG Gln 138	Gln	GAG Glu	GTT Val	TAT Tyr	GCA Ala 1389	Glu	ATT Ile	ATG Met	GCA Ala	GTT Val 139	Leu	AAG Lys	4176
CAT His	GAC Asp	GAT Asp 1395	Gln	CAT His	ACC Thr	ATA Ile	AAT Asn 1400	Thr	CAA Gln	GAC Asp	ATT Ile	GCA Ala 1405	Ser	GAT Asp	CTG Leu	4224
TGT Cys	CAA Gln 1410	Leu	AGT Ser	ACA Thr	CAG Gln	ACT Thr 1415	Val	TTC Phe	TCC Ser	ATG Met	CTT Leu 1420	Asp	CAT His	CTC Leu	ACA Thr	4272
CAG Gln 1429	TGG Trp	GCA Ala	AGG Arg	CAC His	AAA Lys 1430	Phe	CAG Gln	GCA Ala	CTG Leu	AAA Lys 1435	Ala	GAG Glu	AAA Lys	TGT Cys	CCA Pro 1440	4320
CAC His	AGC Ser	AAA Lys	TCA Ser	AAC Asn 1445	Arg	AAT Asn	AAG Lys	GTA Val	GAC Asp 145	Ser	ATG Met	GTA Val	TCT Ser	ACT Thr 1455	Val	4368
GAT Asp	TAT Tyr	GAA Glu	GAC Asp 1460	Tyr	CAG Gln	AGT Ser	GTA Val	ACC Thr 1465	Arg	TTT Phe	CTA Leu	GAC Asp	CTC Leu 1470	Ile	CCC Pro	4416
CAG Gln	GAT Asp	ACT Thr 1475	Leu	GCA Ala	GTA Val	GCT Ala	TCC Ser 1480	Phe	CGC Arg	TCC Ser	AAA Lys	GCA Ala 1485	Tyr	ACA Thr	CGA Arg	4464
GCT Ala	GTA Val 1490	Met	CAC His	TTT Phe	GAA Glu	TCA Ser 1495	Phe	ATT Ile	ACA Thr	GAA Glu	AAG Lys 1500	Lys	CAA Gln	AAT Asn	ATT Ile	4512
CAG Gln 1505	GAA Glu	CAT His	CȚT Leu	GGA Gly	TTT Phe 1510	Leu	CAG Gln	AAA Lys	TTG Leu	TAT Tyr 1515	Ala	GCT Ala	ATG Met	CAT His	GAA Glu 1520	4560
CCT Pro	GAT Asp	GGA Gly	GTG Val	TCC Ser 1525	Gly	GTC Val	AGT Ser	GCA Ala	ATT Ile 1530	Arg	AAG Lys	GCA Ala	GAA Glu	CCA Pro 1535	Ser	4608
CTA Leu	AAA Lys	GAA Glu	CAG Gln 1540	Ile	CTT Leu	GAA Glu	CAT His	GAA Glu 1545	Ser	CTT Leu	GGC Gly	TTG Leu	CTG Leu 1550	Arg	GAT Asp	4656
GCC Ala	ACT Thr	GCT Ala 1555	Cys	TAT Tyr	GAC Asp	AGG Arg	GCT Ala 1560	Ile	CAG Gln	CTA Leu	Glu	CCA Pro 1565	qaA	CAG Gln	ATC Ile	4704
ATT	CAT His 1570	Tyr	CAT His	GGT Gly	GTA Val	GTA Val 1575	Lys	TCC Ser	ATG Met	TTA Leu	GGT Gly 1580	Leu	GGT Gly	CAG Gln	CTG Leu	4752

TCT Ser 1585	Thr	GTT Val	ATC Ile	ACT Thr	CAG Gln 1590	Val	AAT Asn	GGA Gly	GTG Val	CAT His 1595	Ala	AAC Asn	AGG Arg	TCC Ser	GAG Glu 1600	4800
TGG Trp	ACA Thr	GAT Asp	GAA Glu	TTA Leu 1605	Asn	ACG Thr	TAC Tyr	AGA Arg	GTG Val 1610	Glu	GCA Ala	GCT Ala	TGG Trp	AAA Lys 1615	Leu	4848
TCA Ser	CAG Gln	TGG Trp	GAT Asp 1620	Leu	GTG Val	GAA Glu	AAC Asn	TAT Tyr 1625	Leu	GCA Ala	GCA Ala	GAT Asp	GGA Gly 1630	Lys	TCT Ser	4896
ACA Thr	ACA Thr	TGG Trp 1635	AGT Ser	GTC Val	AGA Arg	CTG Leu	GGA Gly 1640	Gln	CTA Leu	TTA Leu	TTA Leu	TCA Ser 1645	Ala	AAA Lys	AAA Lys	4944
AGA Arg	GAT Asp 165	Ile	ACA Thr	GCT Ala	TTT Phe	TAT Tyr 1655	Asp	TCA Ser	CTG Leu	AAA Lys	CTA Leu 1660	Val	AGA Arg	GCA Ala	GAA Glu	4992
CAA Gln 166	Ile	GTA Val	CCT Pro	CTT Leu	TCA Ser 1670	Ala	GCA Ala	AGC Ser	TTT Phe	GAA Glu 1679	Arg	GGC Gly	TCC Ser	TAC Tyr	CAA Gln 1680	5040
CGA Arg	GGA Gly	TAT Tyr	GAA Glu	TAT Tyr 1685	Ile	GTG Val	AGA Arg	TTG Leu	CAC His 1690	Met	TTA Leu	TGT Cys	GAG Glu	TTG Leu 1699	GIU	5088
CAT His	AGC Ser	ATC Ile	AAA Lys 1700	Pro	CTT Leu	TTC Phe	CAG Gln	CAT His 170	Ser	CCA Pro	GGT Gly	GAC Asp	AGT Ser 1710	Ser	CAA Gln	5136
GAA Glu	GAT Asp	TCT Ser 171	CTA Leu 5	AAC Asn	TGG Trp	GTA Val	GCT Ala 172	Arg	CTA Leu	GAA Glu	ATG Met	ACC Thr 172	Gln	AAT Asn	TCC Ser	5184
TAC Tyr	AGA Arg 173	Ala	AAG Lys	GAG Glu	CCT Pro	ATC Ile 173	Leu	GCT Ala	CTC Leu	CGG Arg	AGG Arg 174	Ala	TTA Leu	CTA Leu	AGC Ser	5232
CTC Leu 174	Asn	AAA Lys	AGA Arg	CCA Pro	GAT Asp 175	Tyr	AAT Asn	GAA Glu	ATG Met	GTT Val 175	Gly	GAA Glu	TGC Cys	TGG Trp	CTG Leu 1760	5280
CAG Gln	AGT Ser	GCC Ala	AGG Arg	GTA Val 176	Ala	AGA Arg	AAG Lys	GCT Ala	GGT Gly 177	His	CAC His	CAG Gln	ACA Thr	GCC Ala 177	Tyr	5328
AAT Asn	GCT Ala	CTC Leu	CTT Leu 178	Asn	GCA Ala	GGG Gly	GAA Glu	TCA Ser 178	Arg	CTC Leu	GCT Ala	GAA Glu	CTG Leu 179	Tyr	GTG Val	5376
GAA Glu	AGG Arg	GCA Ala 179	Lys	TGG Trp	CTC Leu	TGG Trp	TCC Ser 180	Lys	GGT Gly	GAT Asp	GTT Val	CAC His 180	Gln	GCA Ala	CTA Leu	5424
ATT Ile	GTT Val 181	Leu	CAA Gln	AAA Lys	GGT Gly	GTT Val 181	Glu	TTA Leu	TGT Cys	TTT Phe	CCT Pro 182	Glu	AAT Asn	GAA Glu	ACC Thr	5472
CCA Pro 182	Pro	GAG Glu	GGT Gly	AAG Lys	AAC Asn 183	Met	TTA Leu	ATC	CAT His	GGT Gly 183	Arg	GCT Ala	ATG Met	CTA Leu	CTA Leu 1840	5520
GTG Val	GGC Gly	CGA Arg	TTT Phe	ATG Met 184	Glu	GAA Glu	ACA Thr	GCT Ala	AAC Asn 185	Phe	GAA Glu	AGC Ser	AAT Asn	GCA Ala 185	ATT Ile 5	5568

ATG Met	AA/ Lys	A AAA	TAT TY1 186	с гла	G GAT	GTC Val	ACC Thr	GCC Ala 186	. Cys	CTG Leu	CCA Pro	A GAA O Glu	TGC Trp 187	Gli	GAT Asp	5,6	16
GGG Gly	CAT His	TTT: Phe	: Туг	CTT Leu	GCC Ala	Lys	TAC Tyr 188	Tyr	GAC Asp	Lys	Leu	ATG Met 188	Pro	ATO Met	GTC Val	56	64
ACA Thr	GAC Asp 189	AST.	AAA Lys	ATG Met	GAA Glu	AAG Lys 189	Gln	GGT Gly	GAT Asp	CTC Leu	ATC Ile 190	Arg	TAT Tyr	'ATA	GTT Val	57	12
CTT Leu 190	HIS	TTT Phe	GGC Gly	AGA Arg	TCT Ser 191	Leu	CAA Gln	TAT Tyr	GGA Gly	AAT Asn 191	Gln	TTC Phe	ATA Ile	TAT Tyr	CAG Gln 1920	57	60
TCA Ser	ATG Met	CCA Pro	CGA Arg	Met 192	Leu	ACT Thr	CTA Leu	TGG Trp	CTT Leu 193	Ąsp	TAT Tyr	GGT Gly	ACA Thr	AAG Lys 193	Ala	586	9.0
TAT Tyr	GAA Glu	TGG Trp	GAA Glu 194	гys	GCT Ala	GGC Gly	CGC Arg	TCC Ser 194	Asp	CGT Arg	GTA Val	CAA Gln	ATG Met 195	Arg	AAT Asn	585	56
GAT Asp	TTG Leu	GGT Gly 195	ьys	ATA	AAC Asn	AAG Lys	GTT Val 1960	Ile	ACA Thr	GAG Glu	CAT	ACA Thr 1965	Asn	тат Туг	TTA Leu	590	04
GCT Ala	CCA Pro 197	Tyr	CAA Gln	TTT Phe	TTG Leu	ACT Thr 1979	Ala	TTT Phe	TCA Ser	CAA Gln	TTG Leu 198	ATC Ile	TCT Ser	CGA Arg	ATT Ile	595	52
TGT Cys 1985	HIS	TCT Ser	CAC His	GAT Asp	GAA Glu 1990	.Val	TTT Phe	GTT Val	GTG Val	CTT Leu 1995	Asp	GGA Gly	AAT Asn	AAT Asn	AGC Ser 2000	600	00
CAA Gln	GTA Val	TTT Phe	CTA Leu	GCC Ala 2009	Tyr	CCT Pro	CAA Gln	CAA Gln	GCA Ala 2010	Met	TGG Trp	ATG Met	ATG Met	ACA Thr 2015	Ala	604	. 8
GTG Val	TCA Ser	AAG Lys	TCA Ser 2020	Ser	TAT Tyr	CCC Pro	ATG Met	CGT Arg 2025	Val	AAC Asn	AGA Arg	TGC Cys	AAG Lys 2030	Glu	ATC Ile	609	6
CTC Leu	AAT Asn	AAA Lys 2039	Ala	ATT Ile	CAT His	ATG Met	AAA Lys 2040	Lys	TCC Ser	TTA Leu	GAG Glu	AAG Lys 2045	Phe	GTT Val	GGA Gly	614	4
Asp	GCA Ala 2050	Inr	CGC Arg	CTA Leu	ACA Thr	GAT Asp 2055	Lys	CTT Leu	CTA Leu	GAA Glu	TTG Leu 2060	TGC Cys	AAT Asn	AAA Lys	CCG Pro	619	2
GTT Val 2065	ASD	GGA Gly	AGT Ser	AGT Ser	TCC Ser 2070	Thr	TTA Leu	AGC Ser	ATG Met	AGC Ser 2075	Thr	CAT His	TTT Phe	AAA Lys	ATG Met 2080	624	0
CTT Leu	AAA Lys	AAG Lys	CTG Leu	GTA Val 2085	Glu	GAA Glu	GCA . Ala	Thr	TTT Phe 2090	Ser	GAA Glu	ATC Ile	CTC Leu	ATT Ile 2095	Pro	628	8
CTA (CAA Gln	TCA Ser	GTC Val 2100	Met	ATA Ile	CCT Pro	Thr :	CTT Leu 2105	CCA Pro	TCA Ser	ATT Ile	Leu	GGT Gly 2110	Thr	CAT His	633	6
GCT A	AAC Asn	CAT His 2115	Ala	AGC Ser	CAT His	Glu	CCA ' Pro : 2120	rrr Phe	CCT Pro	GGA Gly	CAT His	TGG (Trp / 2125	GCC Ala	TAT Tyr	ATT Ile	638	4

ANG ANG ANT TCT TTA ANA GCC TCA GAT GGA ANG TTC TAC ATC ATG ATG LYS LYS Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met 2145 TGT AAG CCA ANA GAT GAC CTG AGA AAG GAT TGT AGA CTA ATG GAA TTC Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe 2175 AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA GAG TCT CGT AGA ASS Ser Leu Ile Ass Lys Cys Leu Arg Lys Asp Asp Ala Glu Ser Arg Arg 2180 AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA AAT GAT GAA ARG GGI Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Ass Asp Glu 2185 TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT TTG AGA CCT ATT CYS Gly Ile Ile Glu Try Val Ass Ass Thr Ala Gly Leu Arg Pro Ile 2215 CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA GAC CTT CAT ATG Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2225 CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2246 CTT CGC CAG TGT ATG CTA CCA AAG TCA GAG GCT TTA TTT CAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 22260 GAG TGG TT CTG AGA ACA TTC CCT GC CC AGG CAT CCT CCT ATT TTT CAT Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2276 GAG TGG TT CTG AGA ACA TTC CCT GAT CCA GG CAT CCT CCT ATT TTT CAT Lys Val Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2275 AAT CTG GGG CTT GGA GAC CTC CAT GCA GTA ATG TCA TTG TTT CAT GGU TTP CAT CTG CCC AGG CAT CCT CCT TAT TTT CAT Lys Val Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2275 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCT TTT GAT TCT GAT ATG ACG GAT CTT CTT GGT CAT ATG ATG GTT ATT CTG GGG CTT GAG ACA TTC CTG CTC ATT TCT TTT GAT TCT LEU GIY Leu Gly Asp Arg His Gly Glu Ass Ile Leu Phe Asp Ser 2330 ATT CTG GGG CTT GGA GAC CTT CAT GTA GAT TCT TTT CAT ATG ACG GAA ACC TTT GAG GAC CTT TTT CAT ATG GLU Thr His 23240 AAT ATG GTT AAT GAG ATG GGT CCT ATG GGA ATT TTT CCC CTG CCT CAT TTT CAT AGG GLU Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340	132
Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe 2175 AAT TCC TTG ATT AAT AAG TGC TTA AGA AAA GAT GCA GAG TCT CGT AGA ASS Ser Leu Ile Ass Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg 2180 AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA AAT GAT GAA ARG GIU Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Ass Asp Glu 2195 TGT GGG ATT ATT GAA TGG GTG ACA ACA CACT GCT GGT TTG AGA CCT ATT Cys Gly Ile Ile Glu Trp Val Ass Ass Thr Ala Gly Leu Arg Pro Ile 2210 CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA GAA CAC ACT GCT GGT TTG AGA CTC ACT Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2240 CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2225 AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT CAT Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 GAG TGG TTT CTG AGA ACA TC CCT GAT CCT ACA TCA TGG TAC AGT AGT GIU Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2227 AGA TCA GCT TAC TGC CGT CCC ACT GCA GTA ATG TCA ATG TGT AGT AGT AGT AGT GIU Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2227 AGA TCA GCT TAC TGC CGT CCC ACT GCA GTA ATG TCA TGT GAT AGT AGT AGT AGT Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGG CTT GGA GAC GTT CTT TCT GAT AGT CTC CTC ACT CTC TTT TCT CTC GGC GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGA GAC CGT CAT GAT GGT GAA AAT ATT CTC TTT GAT TCT CTC GGC GGA GAC CGT CAT GTA GGT TAT TCT GGC CTC ACT CAT CTC CTC GGC GAA ACC TTT GAA GGT TCT TTC CAT GGT GGT GAT AGT TCT CTC GGC GAA CCC TTT GAA GGT TCT TTT CGA CGC GGA GAA ACC TTT GAA GGT CCT ATT GGC CTC ACT CAT GGG GAA AC	80
ASA SET LEU Ile ASA LYS CYS LEU ATG LYS ASP Ala Glu SET ATG ATG 2180 AGA GAA CTT CAT ATT CGA ACA TAT GCA GTT ATT CCA CTA AAT GAT GAA ATG GIL LEU HIS Ile ATG THY TYT Ala Val Ile Pro Leu ASA ASP Glu 2195 TGT GGG ATT ATT GAA TGG GTG ACA AAC ACT GCT GGT TTG AGA CCT ATT CYS Gly Ile Ile Glu Trp Val ASA ASA THY Ala GIL Leu Arg Pro Ile 2210 CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA GAA CAT GCT GCT GAT CYS LEU TYT LYS Glu LYS GILY TYT Met THY GILY LYS LEU TYT LYS GIL LYS GILY TYT MET THY GILY LYS LEU ATG CYS GILY 2220 CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC LEU ARG GIL CYS Met LEU PRO LYS SET Ala Ala LEU SET GIL LYS LEU 2245 AAAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT CAT LYS Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TCA TGT TAT TYP SET SET 2285 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTA GGT TAT ATG ACA GGT TAT ATG ACA GGT AGT TTT CTC CTC CCT GAT CCT ACA TCA TCA TGT TTT CAT GAT AGG TAG TCT CTC ATT TTT CAT GAT AGG TAG TCT CTC ATT TTT CAT GAT AGG TAG TCT CTC ATT TTT CAT GAT AGG TGT TTT CAT GAT AGG TGT TTT TTT CAT GAT AGG TTT TTT CAT GAT AGG TTT TTT CAT GAT AGG TTT TTT TYT SET SET 2285 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT TAT ATG GGT TAT ATG GGT GAT ATG TCT GGG TTT GGA GAC CCT CAT GGT GAA AAT ATT CTC TTT GAT TCT TTT GAT TCT TTT CAT ATG GGT GAT ATG TCT GGG TTT GAT AGG TTT TTT	28
Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu 2295 TGT GGG ATT ATT GAA TGG GTG AAC AAC ACT GCT GGT TTG AGA CCT ATT CYS Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile 2210 CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA GAA CACT GCT GAT THE Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2225 CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT AGT Lys Val Phe Arg Glu Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2280 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA AGT AGT GGT AGT Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGA AAT ATT CTC TTT GAT TCT GAT ARG Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2300 ATT CTG GGG CTT GGA GAC CGT CAT GTA ATT TCT TTT CAT AGG ACT ACT GGG GAC ACT CTT TTT CAT ARG Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2300 ATT CTG GGG CTT GGA GAC CGT CAT GTA ATT TCT TTT CAT AGG TAC ACT CT GGG GAC ACT CTT GAT TCT GAT TCT GAT TCT GAT GGT ACT GAT GGT GAT GAT GTA ATG GTT GAT TCT GAT TCT GAT GAT GAT GAT GAT GAT GAT GAT GAT GA	76
Cys Gly Ile Ile Glu Try Val Asn Asn Thr Ala Gly Leu Arg Pro Ile 2210 CTG ACC AAA CTA TAT AAA GAA AAG GGA GTG TAT ATG ACA GGA AAA GAA Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2225 CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2245 AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT ATT TTT CAT Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT AGT Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2275 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GT Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT Tle Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2310 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2335 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CAT TTT CAT Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	24
Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2230 CTT CGC CAG TGT ATG CTA CCA AAG TCA GCA GCT TTA TCT GAA AAA CTC Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2255 AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT CAT Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro 11e Phe His 2260 GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT AGT Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2275 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT TAT ATG GAT GGT GGT TAT Ala Val Met Ser Met Val Gly Tyr 2295 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT G96 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT 11e Leu Gly Leu Gly Asp Arg His Gly Glu Asn 11e Leu Phe Asp Ser 2310 TTG ACT GGT GAA TGC GTA CAT GAT GAT TTC AAT TGT CTT TC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2335 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT GGT GIU Thr Phe Glu Val Pro Glu 11e Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA ASN Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	72
Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2255 AAA GTA TTC CGA GAA TTT CTC CTG CCC AGG CAT CCT CCT ATT TTT CAT Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT AGT GGI Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2280 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT TAT AGA GSEr Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT GGA GT Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2320 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2330 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT GGI GIU Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA 710 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA 710 AASN Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	20
Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 GAG TGG TTT CTG AGA ACA TTC CCT GAT CCT ACA TCA TGG TAC AGT AGT GIu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2285 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT TAT Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2310 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2325 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	68
Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2285 AGA TCA GCT TAC TGC CGT TCC ACT GCA GTA ATG TCA ATG GTT GGT TAT 691 Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT 11e Leu Gly Leu Gly Asp Arg His Gly Glu Asn 11e Leu Phe Asp Ser 2310 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG 2320 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG 100 Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2335 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT 105 Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Arg Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	16
Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2290 ATT CTG GGG CTT GGA GAC CGT CAT GGT GAA AAT ATT CTC TTT GAT TCT G1e Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2305 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2325 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT G1y Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	64
THE Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2310 TTG ACT GGT GAA TGC GTA CAT GTA GAT TTC AAT TGT CTT TTC AAT AAG Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2325 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	12
Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2335 GGA GAA ACC TTT GAA GTT CCA GAA ATT GTG CCA TTT CGC CTG ACT CAT Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	60
Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2340 2345 AAT ATG GTT AAT GGA ATG GGT CCT ATG GGA ACA GAG GGT CTT TTT CGA Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	80
Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg	56
·	04
AGA GCA TGT GAA GTT ACA ATG AGG CTG ATG CGT GAT CAG CGA GAG CCT 715 Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro 2370 2375 2380	52
TTA ATG AGT GTC TTA AAG ACT TTT CTA CAT GAT CCT CTT GTG GAA TGG 720 Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp 2385 2390 2395 2400	0.0





PCT/U

				Lys	GGG Gly	His	Ser			Pro			_		Gly	7248
				Glu	AAG Lys				His					Glu		7296
			Gly		ATC Ile			Arg					Gly			7344
		Ile			CAT His		His					Glu				7392
	Asn				CAG Gln 2470	Met					Thr					7437
TGA																7440

-88-

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2479 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser 1 5 15 15 Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu 30 Met Ser Met Gln Lys Leu Glu Phe Ile Glu Val Thr Leu Leu Thr Val 35 40 45

Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu 50 55 60

Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile 65 70 75 80

Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly 85 90 95

Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu 100 105 110

Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu 115 120 125

Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu 130 135 140

Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys 145 150 155 160

Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp 165 170 175

Leu	Leu	Lys	Ala 180	Ala	Leu	Cys	His	Leu 185	Leu	Gln	Tyr	Phe	Leu 190	Lys	Phe
Val	Pro	Ala 195	Gly	туг	Glu	Ser	Ala 200	Leu	Gln	Val	Arg	Lys 205	Val	Tyr	Val
Arg	Asn 210	Ile	Сув	Lys	Ala	Leu 215	Leu	Asp	Val	Leu	Gly 220	Ile	Glu	Val	Asp
Ala 225	Glu	Tyr	Leu	Leu	Gly 230	Pro	Leu	Tyr	Ala	Ala 235	Leu	Lys	Met	Glu	Ser 240
Met	Glu	Ile	Ile	Glu 245	Glu	Ile	Gln	Сув	Gln 250	Thr	Gln	Gln	Glu	Asn 255	Leu
Ser	Ser	Asn	Ser 260	qaA	Gly	Ile	Ser	Pro 265	Lys	Arg	Arg	Arg	Leu 270	Ser	Ser
Ser	Leu	Asn 275	Pro	Ser	Lys	Arg	Ala 280	Pro	Lys	Gln	Thr	Glu 285	Glu	Ile	Lys
His	Val 290	Asp	Met	Asn	Gln	Lys 295	Ser	Ile	Leu	Trp	Ser 300	Ala	Leu	Lys	Gln
305					Gln 310					315					320
Pro	Val	Ile	Glu	Met 325	Leu	Glu	Gly	Ile	Ala 330	Val	Val	Leu	Gln	Leu 335	Thr
Ala	Leu	Cys	Thr 340	Val	His	Cys	Ser	His 345	Gln	Asn	Met	Asn	Cys 350	Arg	Thr
Phe	Lys	Asp 355	Cys	Gln	His	Lys	Ser 360	Lys	Lys	Lys	Pro	Ser 365	Val	Val	Ile
Thr	Trp 370	Met	Ser	Leu	Asp	Phe 375	Tyr	Thr	Thr	Val	Leu 380	Lys	Ser	Cys	Arg
Arg 385	Leu	Leu	Glu	Ser	Val 390	Gln	Lys	Arg	Thr	Gly 395	Gly	Asn	Ile	Asp	Lys 400
Val	Val	Lys	Ile	Tyr 405	Asp	Ala	Leu	Ile	Tyr 410	Met	Gln	Val	Asn	Ser 415	Ser
Phe	Glu	Asp	His 420	Ile	Leu	Glu	Asp	Leu 425	Cys	Gly	Met	Leu	Ser 430	Leu	Pro
Trp	Ile	Tyr 435	Ser	His	Ser	Asp	Asp 440	Gl.y	Cys	Leu	Lys	Leu 445	Thr	Thr	Phe
Ala	Ala 450	Asn	Leu	Leu	Thr	Leu 455	Ser	Cys	Arg	Ile	Ser 460	Asp	Ser	Tyr	Ser
Pro 465	Gln	Ala	Gln	Ser	Arg 470	Cys	Val	Phe	Leu	Leu 475	Thr	Leu	Phe	Pro	Arg 480
Arg	Ile	Phe	Leu	Glu 485	Trp	Arg	Thr	Ala	Val 490	Tyr	Asn	Trp	Ala	Leu 495	Gln
Ser	Ser	His	Glu 500	Val	Ile	Arg	Ala	Ser 505	Cys	Val	Ser	Gly	Phe 510	Phe	Ile
Leu	Leu	Gln 515	Gln	Gln	Asn	Ser	Cys 520	Asn	Arg	Val	Pro	Lys 525	Ile	Leu	Ile

Asp	Lys 530	Val	Lys	Asp	Asp	Ser	Asp	o Ile	e Val	l Lys			ı Phe	e Ala	. Se
Ile			, (31-	ı T.o.	ı Mal	535		. 7	, tti -		540		_	_	
545					. 550					555	i		Э Туг		56
				565					570)			Asp	575	i
Сув	Arg	Asn	Leu 580	Lys	Ala	Thr	Ser	Glr 585	His	Glu	Cys	Ser	Ser 590		Glı
		595					600					605			
Ile	Pro 610	Ser	Pro	Val	Lys	Leu 615	Ala	Phe	lle	Asp	Asn 620	Leu	His	His	Let
Cys 625	Lys	His	Leu	Asp	Phe 630	Arg	Glu	Asp	Glu	635	Asp	Val	Lys	Ala	Va] 640
Leu	Gly	Thr	Leu	Leu 645	Asn	Leu	Met	Glu	Asp 650	Pro	Asp	Lys	Asp	Val 655	Arg
Val	Ala	Phe	Ser 660	Gly	Asn	Ile	Lys	His 665	Ile	Leu	Glu	Ser	Leu 670	Asp	Ser
Glu	Asp	Gly 675	Phe	Ile	Lys	Glu	Leu 680	Phe	Val	Leu	Arg	Met 685	Lys	Glu	Ala
Tyr	Thr 690	His	Ala	Gln	Ile	Ser 695	Arg	Asn	Asn	Glu	Leu 700	Lys	Asp	Thr	Leu
Ile 705	Leu	Thr	Thr	Gly	Asp 710	Ile	Gly	Arg	Ala	Ala 715	Lys	Gly	Asp	Leu	Val 720
Pro	Phe	Ala	Leu	Leu 725	His	Leu	Leu	His	Cys 730	Leu	Leu	Ser	Lys	Ser 735	Ala
Ser	Val	Ser	Gly 740	Ala	Ala	Tyr	Thr	Glu 745	Ile	Arg	Ala	Leu	Val 750	Ala	Ala
Lys	Ser	Val 755	Lys	Leu	Gln	Ser	Phe 760	Phe	Ser	Gln	Tyr	Lys 765	Lys	Pro	Ile
Сув	Gln 770	Phe	Leu	Val	Glu	Ser 775	Leu	His	Ser	Ser	Gln 780	Met	Thr	Ala	Leu
Pro 785	Asn	Thr	Pro	Cys	Gln 790	Asn	Ala	Asp	Val	Arg 795	Lys	Gln	Asp	Val	Ala 800
His	Gln	Arg	Glu	Met 805	Ala	Leu	Asn	Thr	Leu 810	Ser	Glu	Ile	Ala	Asn 815	Val
Phe	Asp	Phe	Pro 820	Asp	Leu	Asn	Arg	Phe 825	Leu	Thr	Arg	Thr	Leu 830	Gln	Val
Leu	Leu	Pro 835	qaA	Leu	Ala	Ala	Lys 840	Ala	Ser	Pro	Ala	Ala 845	Ser	Ala	Leu
Ile	Arg 850	Thr	Leu	Gly	Lys	Gln 855	Leu	Asn	Val	Asn	Arg 860	Arg	Glu	Ile	Leu
lle 365	Asn	Asn	Phe	Lys	Tyr 870	Ile	Phe	Ser	His	Leu 875	Val	Cys	Ser	Cys	Ser 880

Lys	Asp	Glu	Leu	Glu 885		·Ala	Leu	His	Туг 890		Lys	Asn	Glu	Thr 895	Glu
Ile	Glu	Leu	Gly 900		Leu	Leu	Arg	Gln 905	Asp	Phe	Gln	Gly	Leu 910		Asn
Glu	Leu	Leu 915		Arg	Ile	Gly	Glu 920	His	Tyr	Gln	Gln	Val 925		Asn	Gly
Leu	Ser 930	Ile	Leu	Ala	Ser	Phe 935	Ala	Ser	Ser	Asp	Asp 940	Pro	Tyr	Gln	Gly
Pro 945	Arg	Asp	Ile	Ile	Ser 950	Pro	Glu	Leu	Met	Ala 955	Asp	Tyr	Leu	Gln	Pro 960
Lys	Leu	Leu	Gly	Ile 965	Leu	Ala	Phe	Phe	Asn 970	Met	Gln	Leu	Leu	Ser 975	Ser
Ser	Val	Gly	Ile 980		Asp	Lys	Lys	Met 985	Ala	Leu	Asn	Ser	Leu 990	Met	Ser
Leu	Met	Lys 995	Leu	Met	Gly	Pro	Lys 100		Val	Ser	Ser	Val 100		Val	Lys
Met	Met 101	Thr 0	Thr	Leu	Arg	Thr 101		Leu	Arg	Phe	Lys 102		Asp	Phe	Pro
Glu 1029	Leu	Суѕ	Cys	Arg	Ala 1030		Asp	Cys	Phe	Val 103		Cys	Leu	Asp	His 1040
Ala	Cys	Leu	Gly	Ser 104	Leu 5	Leu	Ser	His	Val		Val	Ala	Leu	Leu 1055	
Leu	Ile	His	Ile 106		Pro	Lys	Glu	Thr 1069		Ala	Ile	Phe	His 1070		Leu
Ile	Ile	Glu 107		Arg	Asp	Ala	Val 1080		Asp	Phe	Leu	His 108		Ile	Tyr
Phe	Leu 1090	Pro	Asp	His	Pro	Glu 1095		Lys	Lys	Ile	Lys 1100		Val	Leu	Gln
Glu 1105	Tyr	Arg	Lys	Glu	Thr 1110		Glu	Ser	Thr	Asp 1115		Gln	Thr	Thr	Leu 1120
Gln	Leu	Ser	Met	Lys 1125	Ala 5	Ile	Gln	His	Glu 1130		Val	Asp	Val	Arg 1135	
His	Ala	Leu	Thr 1140		Leu	Lys	Glu	Thr 1145		Tyr	Lys	Asn	Gln 1150		Lys
Leu	Ile	Lys 1159	Tyr	Ala	Thr	Asp	Ser 1160		Thr	Val	Glu	Pro 1165		Ile	Ser
Gln	Leu 1170	Val	Thr	Val	Leu	Leu 1175		Gly	Cys	Gln	Asp 1180		Asn	Ser	Gln
Ala 1185	Arg	Leu	Leu	Cys	Gly 1190		Cys	Leu	Gly	Glu 1195		Gly	Ala	Ile	Asp 1200
Pro	Gly	Arg	Leu	Asp 1209	Phe	Ser	Thr	Thr	Glu 1210		Gln	Gly	Lys	Asp 1215	
Thr	Phe	Val	Thr 1220	Gly	Val	Glu	Asp	Ser 1225		Phe	Ala	Tyr	Gly 1230		Leu

- Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg Ala 1235 1240 1245
- Pro Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr Asp 1250 1255 1260
- Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg Arg 1265 1270 1275 1280
- Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr Arg 1285 1290 1295
- Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys Pro 1300 1305 1310
- Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala Ser 1315 1320 1325
- Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser Lys 1330 1335 1340
- Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val Thr 1345 1350 1355
- Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys Asn 1365 1370 1375
- Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu Lys 1380 1385 1390
- His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp Leu 1395 1400 1405
- Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu Thr 1410 1415 1420
- Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys Pro 1425 1430 1435 1440
- His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr Val 1445 1450 1455
- Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile Pro 1460 1465 1470
- Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr Arg 1475 1480 1485
- Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn Ile 1490 1495 1500
- Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His Glu 1505 1510 1515. 1520
- Pro Asp Gly Val Ser Gly Val Ser Ala Ile Arg Lys Ala Glu Pro Ser 1525 1530 1535
- Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg Asp 1540 1545 1550
- Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln Ile 1555 1560 1565
- Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln Leu 1570 1580

WO 97/18323 PCT/US96/19337

-93-

Ser	Thr	Val	Ile	Thr	Gln	Val	Asn	Gly	Val	His	Ala	Asn	Arg	Ser	Glu
1585	5				1590)				1595	5				1600

- Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys Leu 1605 1610 1615
- Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys Ser 1620 1625 1630
- Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Ser Ala Lys Lys 1635 1640 1645
- Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala Glu 1650 1660
- Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr Gln 1665 1670 1675 1680
- Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu Glu 1685 1690 1695
- His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln 1700 1705 1710
- Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser 1715 1720 1725
- Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu Ser 1730 1735 1740
- Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu 1745 1750 1760
- Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr 1765 1770 1775
- Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val 1780 1785 1790
- Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala Leu 1795 1800 1805
- Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu Thr 1810 1815 1820
- Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu Leu 1825 1830 1835 1840
- Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala Ile 1845 1850 1855
- Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu Asp 1860 1865 1870
- Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met Val 1875 1880 1885
- Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile Val 1890 1895 1900
- Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr Gln 1905 1910 1915 1920
- Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys Ala 1925 1930 1935

- Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg Asn 1940 1945 1950
- Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr Leu 1955 1960 1965
- Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg Ile 1970 1975 1980
- Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn Ser 1985 1990 1995 2000
- Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr Ala 2005 2010 2015
- Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu Ile 2020 2025 2030
- Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val Gly
 2035 2040 2045
- Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys Pro 2050 2055 2060
- Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys Met 2065 2070 2075 2080
- Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile Pro 2085 2090 2095
- Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr His 2100 2105 2110
- Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr Ile 2115 2120 2125
- Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys Pro 2130 2135 2140
- Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met 2145 2150 2155 2160
- Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe 2165 2170 2175
- Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg 2180 2185 2190
- Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu 2195 2200 2205
- Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile 2210 2215 2220
- Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu 2225 2230 2235 2240
- Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu 2245 2250 2255
- Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe His 2260 2265 2270
- Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser 2275 2280 2285

WO 97/18323 PCT/US96/19337

-95-

Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr 2295

Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser 2305 2310 2315

Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn Lys 2330

Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr His 2345

Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg 2355 2360

Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro

Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp

Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly 2405 2410

Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln

Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro 2440

Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr Asp 2455

Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met

- (2) INFORMATION FOR SEQ ID NO:19:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA
 - (vii) IMMEDIATE SOURCE:
 - (B) CLONE: Primer oDH26
 - (xi) SEQUENCE DESCRIPTION: SEO ID NO:19:

TGGTTTCTGA GAACATTCCC TGA

(2) INFORMATION FOR SEQ ID NO:20:

- - (A) LENGTH: 9 amino acids
 - (i) SEQUENCE CHARACTERISTICS: (B) TYPE: amino acid(D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (vii) IMMEDIATE SOURCE: (B) CLONE: FLAG tag

BNSDOCID: <WO 9718323A2>

23

WO 97/18323

-96-

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
Met Asp Tyr Lys Asp Asp Asp Lys 1 5	
(2) INFORMATION FOR SEQ ID NO:21:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA	
(vii) IMMEDIATE SOURCE: (B) CLONE: Primer 279-3	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:	
TGGATGATGA CAGCTGTGTC	2 (
(2) INFORMATION FOR SEQ ID NO:22:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA	
(vii) IMMEDIATE SOURCE: (B) CLONE: Primer 279-6	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:	
TGTAGTCGCT GCTCAATGTC	2(
(2) INFORMATION FOR SEQ ID NO:23:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 7624 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: cDNA	
(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 3337562	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:	
CTTGTGAAGA GAATGTTTTA CACTCTTGTT AGTGAAGTTT ATTCTTTAAA AGTCAATCGT	6
CAAGGATTTA GCAAATGAAT TAGCACTTCG GATATACTTG TTTATTTAAT ATCTTTTTTTTTT	2

TTTI	TTT	CAA	AGAA:	TCA	GT A	ATTG	GATC	AT A	ACGA(GACT	TCT	GCGG2	ATT (GCAG	CAACTO	180
CCT	CCTG	ГСА	TTTG	rtac <i>i</i>	AC A	AGAA	AATC:	r GTO	GAAG?	CAT	CTG	rtca:	TTA 1	TAT.	rtctt7	r 240
TTA	AAAG(CAA	GAGT	CCTG	CT A	rrr.	rggg	TAC	CTCA	AAA	AGA	ATTA	TTA (CAACT	TTTT	300
AAG <i>I</i>	ACTT(GT	TTAC	CTCC	A TA	GAAG	TAAA	G TG					GTG Val 5			353
CCA Pro	GTG Val	GTC Val 10	ATG Met	AGC Ser	CGA Arg	TTT Phe	TTA Leu 15	AGT Ser	CAA Gln	TTA Leu	GAT Asp	GAA Glu 20	CAC His	ATG Met	GGA Gly	401
TAT Tyr	TTA Leu 25	CAA Gln	TCA Ser	GCT Ala	CCT Pro	TTG Leu 30	CAG Gln	TTG Leu	ATG Met	AGT Ser	ATG Met 35	CAA Gln	AAT Asn	TTA Leu	GAA Glu	449
			GTC Val													497
GTG Val	TTT Phe	TTT Phe	AGA Arg	AGG Arg 60	CAA Gln	GAA Glu	CTC Leu	TTA Leu	CTT Leu 65	TGG Trp	CAG Gln	ATA Ile	GGT Gly	TGT Cys 70	GTT Val	545
			TAT Tyr 75													593
			CTT Leu													641
			AGC Ser													689
			CAA Gln													737
			TTT Phe													785
GTC Val	TAT Tyr	TTA Leu	AAT Asn 155	ATG Met	CTG Leu	CTG Leu	GAA Glu	AAA Lys 160	CTC Leu	TGT Cys	GTC Val	ATG Met	TTT Phe 165	GAA Glu	GAC Asp	833
			ATG Met													881
			CAG Gln													929
			GTC Val													977
			CTT Leu													1025

CTT	TAT	GCA	GCT	TTG	AAA	ATG	GAA	AGT	ATG	GAA	ATC	TTA	GAG	GAG	ATT	1073
			235					240					245		Ile	
CAA Gln	TGC Cys	CAA Gln 250	Thr	CAA Gln	CAG Gln	GAA Glu	AAC Asn 255	Leu	AGC Ser	AGT Ser	AAT Asn	Ser 260	GAT A sp	GGA Gly	ATA Ile	1121
TCA Ser	CCC Pro 265	Lys	AGG Arg	CGT Arg	CGT Arg	CTC Leu 270	AGC Ser	TCG Ser	TCT Ser	CTA Leu	AAC Asn 275	CCT Pro	TCT Ser	AAA Lys	AGA Arg	1169
GCA Ala 280	Pro	AAA Lys	CAG Gln	ACT Thr	GAG Glu 285	GAA Glu	ATT Ile	AAA Lys	CAT His	GTG Val 290	GAC Asp	ATG Met	AAC Asn	CAA Gln	AAG Lys 295	1217
AGC Ser	ATA Ile	TTA Leu	TGG Trp	AGT Ser 300	GCA Ala	CTG Leu	AAA Lys	CAG Gln	AAA Lys 305	GCT Ala	GAA Glu	TCC Ser	CTT Leu	CAG Gln 310	ATT Ile	1265
TCC Ser	CTT Leu	GAA Glu	TAC Tyr 315	AGT Ser	GGC Gly	CTA Leu	AAG Lys	AAT Asn 320	CCT Pro	GTT Val	ATT Ile	GAG Glu	ATG Met 325	TTA Leu	GAA Glu	1313
GGA Gly	ATT Ile	GCT Ala 330	GTT Val	GTC Val	TTA Leu	CAA Gln	CTG Leu 335	ACT Thr	GCT Ala	CTG Leu	TGT Cys	ACT Thr 340	GTT Val	CAT His	TGT Cys	1361
TCT Ser	CAT His 345	CAA Gln	AAC Asn	ATG Met	AAC Asn	TGC Cys 350	CGT Arg	ACT Thr	TTC Phe	AAG Lys	GAC Asp 355	Cys	CAA Gln	CAT His	AAA Lys	1409
TCC Ser 360	AAG Lys	AAG Lys	AAA Lys	CCT Pro	TCT Ser 365	GTA Val	GTG Val	ATA Ile	ACT Thr	TGG Trp 370	ATG Met	TCA Ser	TTG Leu	GAT Asp	TTT Phe 375	1457
TAC Tyr	ACA Thr	AAA Lys	GTG Val	CTT Leu 380	AAG Lys	AGC Ser	TGT Cys	AGA Arg	AGT Ser 385	TTG Leu	TTA Leu	GAA Glu	TCT Ser	GTT Val 390	CAG Gln	1505
AAA Lys	CTG Leu	GAC Asp	CTG Leu 395	GAG Glu	GCA Ala	ACC Thr	ATT Ile	GAT Asp 400	AAG Lys	GTG Val	GTG Val	AAA Lys	ATT Ile 405	TAT Tyr	GAT Asp	1553
GCT Ala	TTG Leu	ATT Ile 410	TAT Tyr	ATG Met	CAA Gln	GTA Val	AAC Asn 415	AGT Ser	TCA Ser	TTT Phe	GAA Glu	GAT Asp 420	CAT His	ATC Ile	CTG Leu	1601
GAA Glu	GAT Asp 425	TTA Leu	TGT Cys	GGA Gly	ATG Met	CTC Leu 430	TCA Ser	CTT Leu	CCA Pro	TGG Trp	ATT Ile 435	TAT Tyr	TCC Ser	CAT His	TCT Ser	1649
GAT Asp 440	GAT Asp	GGC Gly	TGT Cys	TTA Leu	AAG Lys 445	TTG Leu	ACC Thr	ACA Thr	TTT Phe	GCC Ala 450	GCT Ala	AAT Asn	CTT Leu	CTA Leu	ACA Thr 455	1697
TTA Leu	AGC Ser	TGT Cys	AGG Arg	ATT Ile 460	TCA Ser	GAT Asp	AGC Ser	TAT Tyr	TCA Ser 465	CCA Pro	CAG Gln	GCA Ala	CAA Gln	TCA Ser 470	CGA Arg	1745
TGT Cys	GTG Val	TTT Phe	CTT Leu 475	CTG Leu	ACT Thr	CTG Leu	TTT Phe	CCA Pro 480	AGA Arg	AGA Arg	ATA Ile	TTC Phe	CTT Leu 485	GAG Glu	TGG Trp	1793
AGA Arg	ACA Thr	GCA Ala 490	GTT Val	TAC Tyr	AAC Asn	\mathtt{Trp}	GCC Ala 495	CTG Leu	CAG Gln	AGC Ser	TCC Ser	CAT His 500	GAA Glu	GTA Val	ATC Ile	1841

WO 97/18323 PCT/US96/19337

-99-

						GGA Gly 510										1889
						AAG Lys										1937
						GAA Glu										1985
						TTT Phe										2033
						GTG Val										2081
						TCA Ser 590										2129
						CTG Leu										2177
						CTA Leu										2225
						GTA Val										2273
						AAA Lys										2321
						TCC Ser 670										2369
						ATG Met										2417
						AAG Lys										2465
ATT Ile	GGA Gly	AGG Arg	GCC Ala 715	GCA Ala	AAA Lys	GGA Gly	GAT Asp	TTG Leu 720	GTA Val	CCA Pro	TTT Phe	GCA Ala	CTC Leu 725	TTA Leu	CAC His	2513
						TCC Ser										2561
						CTG Leu 750										2609
AGT Ser 760	TTT Phe	TTC Phe	AGC Ser	CAG Gln	TAT Tyr 765	AAG Lys	AAA Lys	CCC Pro	ATC Ile	TGT Cys 770	CAG Gln	TTT Phe	TTG Leu	GTA Val	GAA Glu 775	2657

ጥርር	Стт	ראר	TCT	አርጥ	CNG	ስጥር). DCD	CCA	CTT	ccc	א א מיי	7 CT	CCT	maa	C. D. C.	
Ser	Leu	His	Ser	Ser 780	Gln	Met	Thr	Ala	Leu 785	Pro	Asn	Thr	Pro	Cys 790	Gln	2705
AAT Asn	GCT Ala	GAC Asp	GTG Val 795	CGA Arg	AAA Lys	CAA Gln	GAT Asp	GTG Val 800	GCT Ala	CAC His	CAG Gln	AGA Arg	GAA Glu 805	ATG Met	GCT Ala	2753
TTA Leu	AAT Asn	ACG Thr 810	TTG Leu	TCT Ser	GAA Glu	ATT Ile	GCC Ala 815	AAC Asn	GTT Val	TTC Phe	GAC Asp	TTT Phe 820	CCT Pro	GAT Asp	CTT Leu	2801
AAT Asn	CGT Arg 825	TTT Phe	CTT Leu	ACT Thr	AGG Arg	ACA Thr 830	TTA Leu	CAA Gln	GTT Val	CTA Leu	CTA Leu 835	CCT Pro	GAT Asp	CTT Leu	GCT Ala	2849
GCC Ala 840	AAA Lys	GCA Ala	AGC Ser	CCT Pro	GCA Ala 845	GCT Ala	TCT Ser	GCT Ala	CTC Leu	ATT Ile 850	CGA Arg	ACT Thr	TTA Leu	GGA Gly	AAA Lys 855	2897
CAA Gln	TTA Leu	AAT Asn	GTC Val	AAT Asn 860	CGT Arg	AGA Arg	GAG Glu	ATT Ile	TTA Leu 865	ATA Ile	AAC Asn	AAC Asn	TTC Phe	AAA Lys 870	TAT Tyr	2945
ATT Ile	TTT Phe	TCT Ser	CAT His 875	TTG Leu	GTC Val	TGT Cys	TCT Ser	TGT Cys 880	TCC Ser	AAA Lys	GAT Asp	GAA Glu	TTA Leu 885	GAA Glu	CGT Arg	2993
GCC Ala	CTT Leu	CAT His 890	TAT Tyr	CTG Leu	AAG Lys	AAT Asn	GAA Glu 895	ACA Thr	GAA Glu	ATT Ile	GAA Glu	CTG Leu 900	GGG Gly	AGC Ser	CTG Leu	3041
TTG Leu	AGA Arg 905	CAA Gln	GAT Asp	TTC Phe	CAA Gln	GGA Gly 910	TTG Leu	CAT His	AAT Asn	GAA Glu	TTA Leu 915	TTG Leu	CTG Leu	CGT Arg	ATT Ile	3089
GGA Gly 920	GAA Glu	CAC His	TAT Tyr	CAA Gln	CAG Gln 925	GTT Val	TTT Phe	AAT Asn	GGT Gly	TTG Leu 930	TCA Ser	ATA Ile	CTT Leu	GCC Ala	TCA Ser 935	3137
TTT Phe	GCA Ala	TCC Ser	AGT Ser	GAT Asp 940	GAT Asp	CCA Pro	TAT Tyr	CAG Gln	GGC Gly 945	CCG Pro	AGA Arg	GAT Asp	ATC Ile	ATA Ile 950	TCA Ser	3185
CCT Pro	GAA Glu	CTG Leu	ATG Met 955	GCT Ala	GAT Asp	TAT Tyr	TTA Leu	CAA Gln 960	CCC Pro	AAA Lys	TTG Leu	TTG Leu	GGC Gly 965	ATT Ile	TTG Leu	3233
GCT Ala	TTT Phe	TTT Phe 970	AAC Asn	ATG Met	CAG Gln	TTA Leu	CTG Leu 975	AGC Ser	TCT Ser	AGT Ser	GTT Val	GGC Gly 980	ATT Ile	GAA Glu	GAT Asp	3281
AAG Lys	AAA Lys 985	ATG Met	GCC Ala	TTG Leu	AAC Asn	AGT Ser 990	TTG Leu	ATG Met	TCT Ser	TTG Leu	ATG Met 995	AAG Lys	TTA Leu	ATG Met	GGA Gly	3329
CCC Pro 1000	Lys	CAT His	GTC Val	AGT Ser	TCT Ser 1005	Val	AGG Arg	GTG Val	AAG Lys	ATG Met 1010	Met	ACC Thr	ACA Thr	CTG Leu	AGA Arg 1015	3377
ACT Thr	GGC Gly	CTT Leu	CGA Arg	TTC Phe 1020	Lys	GAT Asp	GAT Asp	TTT Phe	CCT Pro 1025	Glu	TTG Leu	TGT Cys	TGC Cys	AGA Arg 1030	Ala	 3425
TGG Trp	GAC Asp	TGC Cys	TTT Phe 1035	Val	CGC Arg	TGC Cys	CTG Leu	GAT Asp 1040	His	GCT Ala	TGT Cys	CTG Leu	GGC Gly 1045	Ser	CTT Leu	3473

	AGT Ser		Val					Leu					Ile			3521
	GAA Glu 1065	Thr					His					Glu				3569
	GTG Val					His					Leu					3617
	TTA Leu				Lys					Glu					Thr	3665
	GAG Glu			Asp					Leu					Lys		3713
	CAA Gln		Glu					Arg					Thr			3761
	GAA Glu 1145	Thr					Gln					Lys				3809
	AGT Ser					Pro					Leu			_		3857
	AAA Lys				Asp					Ala					Gly	3905
	TGT Cys			Glu					Asp					Asp		3953
	ACA Thr		Glu					Asp					Thr			4001
	GAT Asp 1225	Ser					Gly					Leu				4049
	CTT Leu)					Asn					Asp					4097
	ATT Ile				Leu					Cys					Thr	4145
	GGC Gly			His					Arg					Val		4193
	ATA Ile		Glu					Thr					Ser			4241
TCA Ser	ACC Thr 1305	Asp	TGG Trp	TCT Ser	GGA Gly	GTA Val 1310	Lys	AAG Lys	CCA Pro	ATT Ile	TAC Tyr 1315	Leu	AGT Ser	AAA Lys	TTG Leu	4289

GGT AGT AAC Gly Ser Asn 1320	TTT GCA GAA Phe Ala Glu 132	Trp Ser Ala	TCT TGG G Ser Trp A 1330	CA GGT TAT La Gly Tyr	CTT ATT Leu Ile 1335	4337
ACA AAG GTT Thr Lys Val	CGA CAT GAT Arg His Asp 1340	CTT GCC AGT Leu Ala Ser	AAA ATT T Lys Ile P 1345	TC ACC TGC The Thr Cys	TGT AGC Cys Ser 1350	4385
ATT ATG ATG	AAG CAT GAT Lys His Asp 1355	TTC AAA GTG Phe Lys Val 136	Thr Ile T	AT CTT CTT Yr Leu Leu 1365	Pro His	4433
ATT CTG GTG Ile Leu Val 137	TAT GTC TTA Tyr Val Leu 0	CTG GGT TGT Leu Gly Cys 1375	AAT CAA G Asn Gln G	AA GAT CAG lu Asp Gln 1380	CAG GAG Gln Glu	4481
GTT TAT GCA Val Tyr Ala 1385	GAA ATT ATG Glu Ile Met	GCA GTT CTA Ala Val Leu 1390	Lys His A	AC GAT CAG sp Asp Gln 395	CAT ACC His Thr	4529
ATA AAT ACC Ile Asn Thr 1400	CAA GAC ATT Gln Asp Ile 140	Ala Ser Asp	CTG TGT C. Leu Cys G 1410	AA CTC AGT ln Leu Ser	ACA CAG Thr Gln 1415	4577
ACT GTG TTC Thr Val Phe	TCC ATG CTT Ser Met Leu 1420	GAC CAT CTC Asp His Leu	ACA CAG TO Thr Gln T: 1425	GG GCA AGG rp Ala Arg	CAC AAA His Lys 1430	4625
TTT CAG GCA Phe Gln Ala	CTG AAA GCT Leu Lys Ala 1435	GAG AAA TGT Glu Lys Cys 144	Pro His S	GC AAA TCA er Lys Ser 1445	Asn Arg	4673
AAT AAG GTA Asn Lys Val 145	GAC TCA ATG Asp Ser Met 0	GTA TCT ACT Val Ser Thr 1455	GTG GAT TO	AT GAA GAC yr Glu Asp 1460	TAT CAG Tyr Gln	4721
AGT GTA ACC Ser Val Thr 1465	CGT TTT CTA Arg Phe Leu	GAC CTC ATA Asp Leu Ile 1470	Pro Gln A	AT ACT CTG sp Thr Leu 475	GCA GTA Ala Val	4769
GCT TCC TTT Ala Ser Phe 1480	CGC TCC AAA Arg Ser Lys 148	Ala Tyr Thr	CGA GCT G Arg Ala Va 1490	TA ATG CAC al Met His	TTT GAA Phe Glu 1495	4817
TCA TTT ATT Ser Phe Ile	ACA GAA AAG Thr Glu Lys 1500	AAG CAA AAT Lys Gln Asn	ATT CAG G Ile Gln G 1505	lu His Leu	GGA TTT Gly Phe 1510	4865
TTA CAG AAA Leu Gln Lys	TTG TAT GCT Leu Tyr Ala 1515	GCT ATG CAT Ala Met His 152	Glu Pro A	AT GGA GTG sp Gly Val 1525	Ala Gly	4913
GTC AGT GCA Val Ser Ala 153	ATT AGA AAG Ile Arg Lys 0	GCA GAA CCA Ala Glu Pro 1535	TCT CTA A	AA GAA CAG ys Glu Gln 1540	ATC CTT Ile Leu	4961
GAA CAT GAA Glu His Glu 1545	AGC CTT GGC Ser Leu Gly	TTG CTG AGG Leu Leu Arg 1550	Asp Ala T	CT GCT TGT hr Ala Cys 555	TAT GAC Tyr Asp	5009
AGG GCT ATT Arg Ala Ile 1560	CAG CTA GAA Gln Leu Glu 156	Pro Asp Gln	ATC ATT C. Ile Ile H. 1570	AT TAC CAT is Tyr His	GGT GTA Gly Val 1575	5057
GTA AAG TCC Val Lys Ser	ATG TTA GGT Met Leu Gly 1580	CTT GGT CAG Leu Gly Gln	CTG TCT AG Leu Ser TI 1585	hr Val Ile	ACT CAG Thr Gln 1590	5105

-103-

GTG AAT GGA GTG CAT Val Asn Gly Val His 1595	: Ala Asn Arg	TCC GAG TGG ACA Ser Glu Trp Thr 1600	GAT GAA TTA AAC 5153 Asp Glu Leu Asn 1605
ACG TAC AGA GTG GAA Thr Tyr Arg Val Glo 1610	GCA GCT TGG Ala Ala Trp 1615	Lys Leu Ser Gln	TGG GAT TTG GTG 5201 Trp Asp Leu Val 1620
GAA AAC TAT TTG GCA Glu Asn Tyr Leu Ala 1625	GCA GAT GGA Ala Asp Gly 1630	AAA TCT ACA ACA Lys Ser Thr Thr 1635	Trp Ser Val Arg
CTG GGA CAG CTA TTA Leu Gly Gln Leu Leu 1640			
TAT GAC TCA CTG AAA Tyr Asp Ser Leu Lys 166	Leu Val Arg	GCA GAA CAA ATT Ala Glu Gln Ile 1665	GTA CCT CTT TCA 5345 Val Pro Leu Ser 1670
GCT GCA AGC TTT GAI Ala Ala Ser Phe Glu 1675	Arg Gly Ser		
GTG AGA TTG CAC ATG Val Arg Leu His Met 1690		Leu Glu His Ser	
TTC CAG CAT TCT CCA Phe Gln His Ser Pro 1705	GGT GAC AGT Gly Asp Ser 1710	TCT CAA GAA GAT Ser Gln Glu Asp 1715	Ser Leu Asn Trp
GTA GCT CGA CTA GAA Val Ala Arg Leu Glu 1720	ATG ACC CAG Met Thr Gln 1725	AAT TCC TAC AGA Asn Ser Tyr Arg 1730	GCC AAG GAG CCT 5537 Ala Lys Glu Pro 1735
ATC CTG GCT CTC CGC Ile Leu Ala Leu Arc 174	Arg Ala Leu		
TAC AAT GAA ATG GTT Tyr Asn Glu Met Val 1755	Gly Glu Cys		
AGA AAG GCT GGT CAC Arg Lys Ala Gly His 1770		Ala Tyr Asn Ala	
GGG GAA TCA CGA CTC Gly Glu Ser Arg Lev 1785	GCT GAA CTG Ala Glu Leu 1790	TAC GTG GAA AGG Tyr Val Glu Arg 1795	Ala Lys Trp Leu
TGG TCC AAG GGT GAT Trp Ser Lys Gly Asg 1800			
GTT GAA TTA TGT TTT Val Glu Leu Cys Phe 182	Pro Glu Asn		
ATG TTA ATC CAT GGT Met Leu Ile His Gly 1835	Arg Ala Met		* = =
GAA ACA GCT AAC TTT Glu Thr Ala Asn Phe 1850		Ala Ile Met Lys	

GTG ACC GCG Val Thr Ala 1865	TGC CTG CCA Cys Leu Pro	GAA TGG G Glu Trp G 1870	GAG GAT GGG Glu Asp Gly	CAT TTT TAC C His Phe Tyr L 1875	TT GCC 5969 eu Ala
AAG TAC TAT Lys Tyr Tyr 1880	GAC AAA TTG Asp Lys Leu 188	Met Pro M	ATG GTC ACA Met Val Thr 1890	GAC AAC AAA A Asp Asn Lys M	TG GAA 6017 et Glu 1895
AAG CAA GGT Lys Gln Gly	GAT CTC ATC Asp Leu Ile 1900	CGG TAT A	TA GTT CTT le Val Leu 1905	CAT TTT GGC A His Phe Gly A 1	GA TCT 6065 rg Ser 910
CTA CAA TAT Leu Gln Tyr	GGA AAT CAG Gly Asn Gln 1915	Phe Ile T	YAT CAG TCA Yr Gln Ser 1920	ATG CCA CGA A Met Pro Arg M 1925	TG TTA 6113 et Leu
ACT CTA TGG Thr Leu Trp 193	Leu Asp Tyr	GGT ACA AGGLY Thr Ly 1935	AG GCA TAT	GAA TGG GAA A Glu Trp Glu L 1940	AA GCT 6161 ys Ala
GGC CGC TCC Gly Arg Ser 1945	GAT CGT GTA Asp Arg Val	CAA ATG AG Gln Met A: 1950	GG AAT GAT rg Asn Asp	TTG GGT AAA A Leu Gly Lys I 1955	TA AAC 6209 le Asn
AAG GTT ATC Lys Val Ile 1960	ACA GAG CAT Thr Glu His 1965	Thr Asn Ty	AT TTA GCT Tyr Leu Ala 1970	CCA TAT CAA T Pro Tyr Gln P	TT TTG 6257 ne Leu 1975
ACT GCT TTT Thr Ala Phe	TCA CAA TTG Ser Gln Leu 1980	ATC TCT CO	GA ATT TGT rg Ile Cys 1985	CAT TCT CAC G His Ser His A	AT GAA 6305 sp Glu 990
		Gly Asn As		GTA TTT CTA G Val Phe Leu A 2005	
CCT CAA CAA Pro Gln Gln 2010	Ala Met Trp	ATG ATG AC Met Met Th 2015	CA GCT GTG hr Ala Val	TCA AAG TCA TO Ser Lys Ser So 2020	CT TAT 6401 er Tyr
CCC ATG CGT Pro Met Arg 2025	GTG AAC AGA Val Asn Arg	TGC AAG GA Cys Lys GI 2030	lu Ile Leu	AAT AAA GCT A' Asn Lys Ala I' 2035	TT CAT 6449 Le His
ATG AAA AAA Met Lys Lys 2040	TCC TTA GAG Ser Leu Glu 2045	Lys Phe Va	TT GGA GAT al Gly Asp 2050	GCA ACT CGC C Ala Thr Arg Le	TA ACA 6497 eu Thr 2055
GAT AAG CTT Asp Lys Leu	CTA GAA TTG Leu Glu Leu 2060	TGC AAT AA Cys Asn Ly	AA CCG GTG ys Pro Val 2065	GAA ATT CTT G Glu Ile Leu A 2	CT TCT 6545 la Ser 070
CTT CAG AAA Leu Gln Lys	CCA AAG AAG Pro Lys Lys 2075	Ile Ser Le	TA AAA GGC eu Lys Gly 080	TCA GAT GGA A Ser Asp Gly Ly 2085	AG TTC 6593 vs Phe
TAC ATC ATG Tyr Ile Met 2090	Met Cys Lys	CCA AAA GA Pro Lys As 2095	AT GAC CTG . sp Asp Leu .	AGA AAG GAT TO Arg Lys Asp Cy 2100	GT AGA 6641 /s Arg
CTA ATG GAA Leu Met Glu 2105	TTC AAT TCC Phe Asn Ser	TTG ATT AZ Leu Ile As 2110	sn Lys Cys	TTA AGA AAA G Leu Arg Lys A 2115	AT GCA 6689 Sp Ala
GAG TCT CGT Glu Ser Arg 2120	AGA AGA GAA Arg Arg Glu 2125	Leu His I	TT CGA ACA le Arg Thr 2130	TAT GCA GTT A Tyr Ala Val I	TT CCA 6737 Le Pro 2135

					Gly					Val					GGT Gly 0	6785
				Leu					Lys				GTG Val 216	Tyr		6833
			Glu					Met					GCA Ala			6881
		Lys					Arg					Pro	AGG Arg			6929
	Ile					Phe					Pro		CCT Pro			6977
					Ser					Ser			GTA Val		Ser	7025
				Ile					Asp				GAA Glu 2245	Asn		7073.
			Ser					Cys					TTC Phe			7121
		Asn					Phe					Ile	GTG Val			7169
	Leu					Val					Pro		GGA Gly			7217
					Ala					Met			ATG Met		Asp	7265
				Leu					Lys				CAT His 2325	Asp		7313
			Trp					Lys					GCG Ala			7361
AAT Asn	GAA Glu 2345	Thr	GGA Gly	GAA Glu	GTT Val	GTC Val 2350	Asn	GAA Glu	AAG Lys	GCC Ala	AAG Lys 2355	Thr	CAT His	GTT Val	CTT Leu	7409
	Ile					Gln					Thr		AAT Asn			7457
ACA Thr	GGA Gly	CTG Leu	CCG Pro	TTA Leu 2380	Ser	ATT Ile	GAA Glu	GGA Gly	CAT His 2385	Val	CAT His	TAC Tyr	CTT Leu	ATA Ile 2390	Gln	7505
				Glu					Gln				GGT Gly 2405	Trp		7553

WO 97/18323

-106-

CCA TAT ATG TGAAATGAAA TTATGTAAAA GAATATGTTA ATAATCTAAA Pro Tyr Met 2410

7602

AGTAAAAAA AAAAAAAAA AA

7624

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2410 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val 35 40 45

Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu 50 55 60

Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile
65 70 75 80

Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly 85 90 95

Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu 100 105 110

Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu 115 120 125

Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu 130 135 140

Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys 145 150 155 160

Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp 165 170 175

Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe 180 185 190

Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val

Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp 210 215 220

Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser 225 230 235 240

Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Glu Asn Leu 245 250 255 -107-

Ser	Ser	Asn	Ser 260	Asp	Gly	Ile	Ser	Pro 265	Lys	Arg	Arg	Arg	Leu 270	Ser	Ser
Ser	Leu	Asn 275	Pro	Ser	Lys	Arg	Ala 280	Pro	Lys	Gln	Thr	Glu 285	Glu	Ile	Lys
His	Val 290	Asp	Met	Asn	Gln	Lys 295	Ser	Ile	Leu	Trp	Ser 300	Ala	Leu	Lys	Gln
Lys 305	Ala	Glu	Ser	Leu	Gln 310	Ile	Ser	Leu	Glu	Tyr 315	Ser	Gly	Leu	Lys	Asn 320
Pro	Val	Ile	Glu	Met 325	Leu	Glu	Gly	Ile	Ala 330	Val	Val	Leu	Gln	Leu 335	Thr
Ala	Leu	Суѕ	Thr 340	Val	His	Cys	Ser	His 345	Gln	Asn	Met	Asn	Cys 350	Arg	Thr
Phe	Lys	Asp 355	Cys	Gln	His	Lys	Ser 360	Lys	Lys	Lys	Pro	Ser 365	Val	Val	Ile
Thr	Trp 370	Met	Ser	Leu	Asp	Phe 375	Tyr	Thr	Lys	Val	Leu 380	Lys	Ser	Cys	Arg
Ser 385	Leu	Leu	Glu	Ser	Val 390	Gln	Lys	Leu	Asp	Leu 395	Glu	Ala	Thr	Ile	Asp 400
Lys	Val	Val	Lys	Ile 405	Tyr	Asp	Ala	Leu	Ile 410	Tyr	Met	Gln	Val	Asn 415	Ser
Ser	Phe	Glu	Asp 420	His	Ile	Leu	Glu	Asp 425	Leu	Cys	Gly	Met	Leu 430	Ser	Leu
Pro	Trp	Ile 435	Tyr	Ser	His	Ser	Asp 440	Asp	Gly	Cys	Leu	Lys 445	Leu	Thr	Thr
Phe	Ala 450	Ala	Asn	Leu	Leu	Thr 455	Leu	Ser	Cys	Arg	Ile 460	Ser	Asp	Ser	Tyr
Ser 465	Pro	Gln	Ala	Gln	Ser 470	Arg	Cys	Val	Phe	Leu 475	Leu	Thr	Leu	Phe	Pro 480
Arg	Arg	Ile	Phe	Leu 485	Glu	Trp	Arg	Thr	Ala 490	Val	Tyr	Asn	Trp	Ala 495	Leu
Gln	Ser	Ser	His 500	Glu	Val	Ile	Arg	Ala 505	Ser	Cys	Val	Ser	Gly 510	Phe	Phe
Ile	Leu	Leu 515	Gln	Gln	Gln	Asn	Ser 520	Cys	Asn	Arg	Val	Pro 525	Lys	Ile	Leu
Ile	Asp 530	Lys	Val	Lys	Asp	Asp 535	Ser	Asp	Ile	Val	Lys 540	Lys	Glu	Phe	Ala
Ser 545	Ile	Leu	Gly	Gln	Leu 550	Val	Суѕ	Thr	Leu	His 555	Gly	Met	Phe	Tyr	Leu 560
Thr	Ser	Ser	Leu	Thr 565	Glu	Pro	Phe	Ser	Glu 570	His	Gly	His	Val	Asp 575	Leu
Phe	Ċys	Arg	Asn 580	Leu	Lys	Ala	Thr	Ser 585	Gln	His	Glu	Cys	Ser 590	Ser	Ser
Gln	Leu	Lys 595	Ala	Ser	Val	Cys	Lys 600	Pro	Phe	Leu	Phe	Leu 605	Leu	Lys	Lys

Lys	610	Pro	Ser	Pro	Val	Lys 615	Leu	Ala	Phe	Ile	Asp 620	Asn	Leu	His	His
Let 625	Cys	Lys	His	Leu	Asp 630	Phe	Arg	Glu	Asp	Glu 635	Thr	Asp	Val	Lys	Ala 640
Val	Leu	Gly	Thr	Leu 645	Leu	Asn	Leu	Met	Glu 650	Asp	Pro	Asp	Lys	Asp 655	Val
Arg	Val	Ala	Phe 660	Ser	Gly	Asn	Ile	Lys 665	His	Ile	Leu	Glu	Ser 670	Leu	Asp
Ser	Glu	Asp 675	Gly	Phe	Ile	Lys	Glu 680	Leu	Phe	Val	Leu	Arg 685	Met	Lys	Glu
Ala	Tyr 690	Thr	His	Ala	Gln	Ile 695	Ser	Arg	Asn	Asn	Glu 700	Leu	Lys	Asp	Thr
Leu 705	·Ile	Leu	Thr	Thr	Gly 710	Asp	Ile	Gly	Arg	Ala 715	Ala	Lys	Gly	Asp	Leu 720
Val	Pro	Phe	Ala	Leu 725	Leu	His	Leu	Leu	His 730	Cys	Leu	Leu	Ser	Lys 735	Ser
Ala	Ser	Val	Ser 740	Gly	Ala	Ala	Tyr	Thr 745	Glu	Ile	Arg	Ala	Leu 750	Val	Ala
Ala	Lys	Ser 755	Val	Lys	Leu	Gln	Ser 760	Phe	Phe	Ser	Gln	Tyr 765	Lys	Lys	Pro
Ile	770	Gln	Phe	Leu	Val	Glu 775	Ser	Leu	His	Ser	Ser 780	Gln	Met	Thr	Ala
L eu 785	Pro	Asn	Thr	Pro	Cys 790	Gln	Asn	Ala	Asp	Val 795	Arg	Lys	Gln	Asp	Val 800
Ala	His	Gln	Arg	Glu 805	Met	Ala	Leu	Asn	Thr 810	Leu	Ser	Glu	Ile	Ala 815	Asn
Val	Phe		Phe 820	Pro	Asp	Leu	Asn	Arg 825	Phe	Leu	Thr	Arg	Thr 830	Leu	Gln
Val	Leu	Leu 835	Pro	Asp	Leu	Ala	Ala 840	Lys	Ala	Ser	Pro	Ala 845	Ala	Ser	Ala
Let	Ile 850	Arg	Thr	Leu	Gly	Lys 855	Gln	Leu	Asn	Val	Asn 860	Arg	Arg	Glu	Ile
Let 865	ılle	Asn	Asn	Phe	Lys 870	Tyr	Ile	Phe	Ser	His 875	Leu	Val	Cys	Ser	Cys 880
Ser	Lys	Asp	Glu	Leu 885	Glu	Arg	Ala	Leu	His 890	Tyr	Leu	Lys	Asn	Glu 895	Thr
Glı	ılle	Glu	Leu 900	Gly	Ser	Leu	Leu	Arg 905		Asp	Phe	Gln	Gly 910	Leu	His
Asr	Glu	Leu 915	Leu	Leu	Arg	Ile	Gly 920	Glu	His	Tyr	Gln	Gln 925	Val	Phe	Asn
Gly	930	Ser	Ile	Leu	Ala	Ser 935	Phe	Ala	Ser	Ser	Asp 940	Asp	Pro	Tyr	Gln
Gly 945	Pro	Arg	Asp	Ile	Ile 950	Ser	Pro	Glu	Leu	Met 955	Ala	Asp	Tyr	Leu	Gln 960

-109-

Pro	Lys	Leu	Leu	Gly 965	Ile	Leu	Ala	Phe	Phe 970	Asn	Met	Gln	Leu	Leu 975	Ser
Ser	Ser	Val	Gly 980	Ile	Glu	Asp	Lys	Lys 985	Met	Ala	Leu	Asn	Ser 990	Leu	Met
Ser	Leu	Met 995	Lys	Leu	Met	Gly	Pro 1000	Lys	His	Val	Ser	Ser 1005	Val	Arg	Val
Lys	Met 1010	Met	Thr	Thr	Leu	Arg 1015		Gly	Leu	Arg	Phe 1020		Asp	Asp	Phe
Pro 1025		Leu	Cys	Cys	Arg 1030		Trp	Asp	Cys	Phe 1035		Arg	Cys	Leu	Asp 1040
His	Ala	Cys	Leu	Gly 1045		Leu	Leu	Ser	His 1050		Ile	Val	Ala	Leu 1055	
Pro	Leu	Ile	His 1060		Gln	Pro	Lys	Glu 1065		Ala	Ala	Ile	Phe 1070		Tyr
Leu	Ile	Ile 1075		Asn	Arg	Asp	Ala 1080	Val	Gln	Asp	Phe	Leu 1085	His	Glu	Ile
Tyr	Phe 1090	Leu)	Pro	Asp	His	Pro 1095		Leu	Lys	Lys	Ile 1100		Ala	Val	Leu
Gln 1105		Tyr	Arg	Lys	Glu 1110		Ser	Glu	Ser	Thr 1115		Leu	Gln	Thr	Thr 1120
Leu	Gln	Leu	Ser	Met 1125		Ala	Ile	Gln	His 1130		Asn	Val	Asp	Val 1135	
Ile	His	Ala	Leu 1140		Ser	Leu	Lys	Glu 1145		Leu	Tyr	Lys	Asn 1150		Glu
Lys	Leu	Ile 1159		Tyr	Ala	Thr	Asp 1160		Glu	Thr	Val	Glu 1165		Ile	Ile
Ser	Gln 1170	Leu)	Val	Thr	Val	Leu 1175		Lys	Gly	Cys	Gin 1180		Ala	Asn	Ser
Gln 1189		Arg	Leu	Leu	Cys 1190		Glu	Cys	Leu			Leu	Gly	Ala	Ile 1200
Asp	Pro	C1								1195	•				
		GIÀ	Arg	Leu 1209		Phe	Ser	Thr	Thr 1210	Glu		Gln	Gly	Lys 1215	
Phe	Thr	Phe	_	1205 Thr	5				1210 Ser	Glu)	Thr			1215 Gly	5
		_	Val 1220 Leu	1205 Thr	Gly	Val	Glu	Asp 1225 Leu	1210 Ser	Glu) Ser	Thr Phe	Ala	Tyr 1230 Asn	1215 Gly)	Leu
Leu	Met	Phe Glu 1235 Asp	Val 1220 Leu	1209 Thr) Thr	Gly Arg	Val Ala	Glu Tyr 1240 Ala	Asp 1225 Leu	Ser Ala	Glu Ser Tyr	Thr Phe Ala	Ala Asp 1245 Leu	Tyr 1230 Asn	Gly Ser	Leu Arg
Leu Ala	Met Gln 1256 Cys	Phe Glu 1235 Asp	Val 1220 Leu Ser	Thr) Thr Ala	Gly Arg	Val Ala Tyr 1259	Glu Tyr 1240 Ala	Asp 1225 Leu	Ser Ala Gln	Glu Ser Tyr Glu	Thr Phe Ala Leu 1260	Ala Asp 1245 Leu	Tyr 1230 Asn Ser	Gly Ser	Leu Arg Tyr
Leu Ala Asp 126	Met Gln 1250 Cys	Phe Glu 1235 Asp	Val 1220 Leu Ser	Thr Thr Ala Met	Gly Arg Ala Glu 1270 Val	Val Ala Tyr 1255 Thr	Glu Tyr 1240 Ala S	Asp 1225 Leu Ile Gly	Ser Ala Gln Pro	Glu Ser Tyr Glu Gly 1275	Thr Phe Ala Leu 1260	Ala Asp 1245 Leu)	Tyr 1230 Asn Ser Leu	Gly Ser Ile	Leu Arg Tyr Arg 1280

- Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala 1315 1320 1325
- Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1330 1340
- Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val 1345 1350 1355 1360
- Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1365 1370 1375
- Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 1385 1390
- Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395 1400 1405
- Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 1415 1420
- Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 1430 1435 1440
- Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 1450 1455
- Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 1465 1470
- Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 1480 1485
- Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 1495 1500
- Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 1510 1515
- Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 1530 1535
- Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 1545 1550
- Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 1560 1565
- Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1570 1575 1580
- Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser 1585 1590 1595 1600
- Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys 1605 1610 1615
- Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys 1620 1625 1630
- Ser Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Leu Ser Ala Lys 1635 1640 1645
- Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala 1650 1655 1660

WO 97/18323 PCT/US96/19337

-111-

Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr 1670 1675 Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu 1685 1690 Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu 1730 1735 1740 Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala 1800 Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu 1815 Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala 1850 Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu 1865 Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met 1880 Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile 1895 1900 Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr 1910 1915 Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys 1925 1930 Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg 1945 Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr 1955 1960 Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg 1970 1975 1980 Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn 1990 1995

Ser Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr

2010

2005

- Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu 2020 2025 2030
- Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val 2035 2040 2045
- Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys 2050 2055 2060
- Pro Val Glu Ile Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu 2065 2070 2075 2080
- Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp 2085 2090 2095
- Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn 2100 2105 2110
- Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg Arg Glu Leu His Ile 2115 2120 2125
- Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu 2130 2135 2140
- Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr 2145 2150 2155 2160
- Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met 2165 2170 2175
- Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu 2180 2185 2190
- Phe Leu Pro Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg 2195 2200 2205
- Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys 2210 2215 2220
- Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly 2225 2230 2235 2240
- Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys 2245 2250 2255
- Val His Val Asp Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu 2260 2265 2270
- Val Pro Glu Ile Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly 2275 2280 2285
- Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val 2290 2300
- Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu 2305 2310 2315 2320
- Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys 2325 2330 2335
- Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu 2340 2345 2350
- Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val 2355 2360 2365

-113-

Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly 2375 2380

His Val His Tyr Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys 2385 2390 2395 2400

Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met 2405

- (2) INFORMATION FOR SEQ ID NO:25:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7502 base pairs
 - (B) TYPE: nucleic acid (C) STRANDEDNESS: single

 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..7440
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATG Met 1	GGT Gly	CAT His	GCT Ala	GTG Val 5	GAA Glu	TGG Trp	CCA Pro	GTG Val	GTC Val 10	ATG Met	AGC Ser	CGA Arg	TTT Phe	TTA Leu 15	AGT Ser	48
CAA Gln	TTA Leu	GAT Asp	GAA Glu 20	CAC His	ATG Met	GGA Gly	TAT Tyr	TTA Leu 25	CAA Gln	TCA Ser	GCT Ala	CCT Pro	TTG Leu 30	CAG Gln	TTG Leu	96
ATG Met	AGT Ser	ATG Met 35	CAA Gln	AAT Asn	TTA Leu	GAA Glu	TTT Phe 40	ATT Ile	GAA Glu	GTC Val	ACT Thr	TTA Leu 45	TTA Leu	ATG Met	GTT Val	144
CTT Leu	ACT Thr 50	CGT Arg	ATT Ile	ATT Ile	GCA Ala	ATT Ile 55	GTG Val	TTT Phe	TTT Phe	AGA Arg	AGG Arg 60	CAA Gln	GAA Glu	CTC Leu	TTA Leu	192
CTT Leu 65	TGG Trp	CAG Gln	ATA Ile	GGT Gly	TGT Cys 70	GTT Val	CTG Leu	CTA Leu	GAG Glu	TAT Tyr 75	GGT Gly	AGT Ser	CCA Pro	AAA Lys	ATT Ile 80	240
AAA Lys	TCC Ser	CTA Leu	GCA Ala	ATT Ile 85	AGC Ser	TTT Phe	TTA Leu	ACA Thr	GAA Glu 90	CTT Leu	TTT Phe	CAG Gln	CTT Leu	GGA Gly 95	GGA Gly	288
CTA Leu	CCA Pro	GCA Ala	CAA Gln 100	CCA Pro	GCT Ala	AGC Ser	ACT Thr	TTT Phe 105	TTC Phe	AGC Ser	TCA Ser	TTT Phe	TTG Leu 110	GAA Glu	TTA Leu	336
TTA Leu	AAA Lys	CAC His 115	CTT Leu	GTA Val	GAA Glu	ATG Met	GAT Asp 120	ACT Thr	GAC Asp	CAA Gln	TTG Leu	AAA Lys 125	CTC	TAT Tyr	GAA Glu	384
GAG Glu	CCA Pro 130	TTA Leu	TCA Ser	AAG Lys	CTG Leu	ATA Ile 135	AAG Lys	ACA Thr	CTA Leu	TTT Phe	CCC Pro 140	TTT Phe	GAA Glu	GCA Ala	GAA Glu	432
GCT Ala 145	TAT Tyr	AGA Arg	AAT Asn	ATT Ile	GAA Glu 150	CCT Pro	GTC Val	TAT Tyr	TTA Leu	AAT Asn 155	ATG Met	CTG Leu	CTG Leu	GAA Glu	AAA Lys 160	480

CTC Leu	TGT Cys	GTC Val	ATG Met	TTT Phe 165	GAA Glu	GAC Asp	GGT Gly	GTG Val	CTC Leu 170	ATG Met	CGG Arg	CTT Leu	AAG Lys	TCT Ser 175	GAT Asp	528
TTG Leu	CTA Leu	AAA Lys	GCA Ala 180	GCT Ala	TTG Leu	TGC Cys	CAT His	TTA Leu 185	CTG Leu	CAG Gln	TAT Tyr	TTC Phe	CTT Leu 190	AAA Lys	TTT Phe	576
GTG Val	CCA Pro	GCT Ala 195	GGG Gly	TAT Tyr	GAA Glu	TCT	GCT Ala 200	TTA Leu	CAA Gln	GTC Val	AGG Arg	AAG Lys 205	GTC Val	TAT Tyr	GTG Val	624
AGA Arg	AAT Asn 210	ATT Ile	TGT Cys	AAA Lys	GCT Ala	CTT Leu 215	TTG Leu	GAT Asp	GTG Val	CTT Leu	GGA Gly 220	ATT Ile	GAG Glu	GTA Val	GAT Asp	672
GCA Ala 225	GAG Glu	TAC Tyr	TTG Leu	TTG Leu	GGC Gly 230	CCA Pro	CTT	TAT Tyr	GCA Ala	GCT Ala 235	TTG Leu	AAA Lys	ATG Met	GAA Glu	AGT Ser 240	720
ATG Met	GAA Glu	ATC Ile	ATT Ile	GAG Glu 245	GAG Glu	ATT Ile	CAA Gln	TGC Cys	CAA Gln 250	ACT Thr	CAA Gln	CAG Gln	GAA Glu	AAC Asn 255	CTC Leu	768
AGC Ser	AGT Ser	AAT Asn	AGT Ser 260	GAT Asp	GGA Gly	ATA Ile	TCA Ser	CCC Pro 265	AAA Lys	AGG Arg	CGT Arg	CGT Arg	CTC Leu 270	AGC Ser	TCG Ser	816
TCT Ser	CTA Leu	AAC Asn 275	CCT Pro	TCT Ser	AAA Lys	AGA Arg	GCA Ala 280	CCA Pro	AAA Lys	CAG Gln	ACT Thr	GAG Glu 285	GAA Glu	ATT Ile	AAA Lys	864
CAT His	GTG Val 290	GAC Asp	ATG Met	AAC Asn	CAA Gln	AAG Lys 295	AGC Ser	ATA Ile	TTA Leu	TGG Trp	AGT Ser 300	GCA Ala	CTG Leu	AAA Lys	CAG Gln	912
AAA Lys 305	GCT Ala	GAA Glu	TCC Ser	CTT Leu	CAG Gln 310	ATT Ile	TCC Ser	CTT Leu	GAA Glu	TAC Tyr 315	AGT Ser	GGC Gly	CTA Leu	AAG Lys	AAT Asn 320	960
CCT Pro	GTT Val	ATT	GAG Glu	ATG Met 325	TTA Leu	GAA Glu	GGA Gly	ATT Ile	GCT Ala 330	GTT Val	GTC Val	TTA Leu	CAA Gln	CTG Leu 335	ACT Thr	1008
GCT Ala	CTG Leu	TGT Cys	ACT Thr 340	GTT Val	CAT His	TGT Cys	TCT Ser	CAT His 345	CAA Gln	AAC Asn	ATG Met	AAC Asn	TGC Cys 350	CGT Arg	ACT Thr	1056
TTC Phe	AAG Lys	GAC Asp 355	T GT Cys	CAA Gln	CAT His	AAA Lys	TCC Ser 360	AAG Lys	AAG Lys	AAA Lys	CCT Pro	TCT Ser 365	GTA Val	GTG Val	ATA Ile	1104
ACT Thr	TGG Trp 370	ATG Met	TCA Ser	TTG Leu	GAT Asp	TTT Phe 375	TAC Tyr	ACA Thr	AAA Lys	GTG Val	CTT Leu 380	AAG Lys	AGC Ser	TGT Cys	AGA Arg	1152
AGT Ser 385	TTG Leu	TTA Leu	GAA Glu	TCT Ser	GTT Val 390	CAG Gln	AAA Lys	CTG Leu	GAC Asp	CTG Leu 395	GAG Glu	GCA Ala	ACC Thr	ATT Ile	GAT Asp 400	1200
AAG Lys	GTG Val	GTG Val	AAA Lys	ATT Ile 405	TAT Tyr	GAT Asp	GCT Ala	TTG Leu	ATT Ile 410	TAT Tyr	ATG Met	CAA Gln	GTA Val	AAC Asn 415	AGT Ser	1248
TCA Ser	TTT Phe	GAA Glu	GAT Asp 420	CAT His	ATC Ile	CTG Leu	GAA Glu	GAT Asp 425	TTA Leu	TGT Cys	GGA Gly	ATG Met	CTC Leu 430	TCA Ser	CTT Leu	1296

			Tyr					Asp							ACA Thr	134
TTT Phe	GCC Ala 450	Ala	AAT Asn	CTT Leu	CTA Leu	ACA Thr 455	TTA Leu	AGC Ser	TG T Cys	AGG Arg	ATT Ile 460	Ser	GAT Asp	AGC Ser	TAT Tyr	139
	Pro				TCA Ser 470											144
AGA Arg	AGA Arg	ATA Ile	TTC Phe	CTT Leu 485	GAG Glu	TGG Trp	AGA Arg	ACA Thr	GCA Ala 490	GTT Val	TAC Tyr	AAC Asn	TGG Trp	GCC Ala 495	CTG Leu	148
CAG Gln	AGC Ser	TCC Ser	CAT His 500	GAA Glu	GTA Val	ATC Ile	CGG Arg	GCT Ala 505	AGT Ser	TGT Cys	GTT Val	AGT Ser	GGA Gly 510	TTT Phe	TTT Phe	153
ATC Ile	TTA Leu	TTG Leu 515	CAG Gln	CAG Gln	CAG Gln	AAT Asn	TCT Ser 520	TGT Cys	AAC Asn	AGA Arg	GTT Val	CCC Pro 525	AAG Lys	ATT Ile	CTT Leu	158
					GAT Asp											163
TCT Ser 545	ATA Ile	CTT Leu	GGT Gly	CAA Gln	CTT Leu 550	GTC Val	TGT Cys	ACT Thr	CTT Leu	CAC His 555	GGC Gly	ATG Met	TTT Phe	TAT Tyr	CTG Leu 560	168
ACA Thr	AGT Ser	TCT Ser	TTA Leu	ACA Thr 565	GAA Glu	CCT Pro	TTC Phe	TCT Ser	GAA Glu 570	CAC His	GGA Gly	CAT His	GTG Val	GAC Asp 575	CTC Leu	172
					AAA Lys											177
					GTC Val											182
AAA Lys	ATA Ile 610	CCT Pro	AGT Ser	CCA Pro	GTA Val	AAA Lys 615	CTT Leu	GCT Ala	TTC Phe	ATA Ile	GAT Asp 620	TAA Asn	CTA Leu	CAT His	CAT. His	187
CTT Leu 625	TGT Cys	AAG Lys	CAT His	CTT Leu	GAT Asp 630	TTT Phe	AGA Arg	GAA Glu	GAT Asp	GAA Glu 635	ACA Thr	GAT Asp	GTA Val	AAA Lys	GCA Ala 640	1920
GTT Val	CTT Leu	GGA Gly	ACT Thr	TTA Leu 645	TTA Leu	AAT Asn	TTA Leu	ATG Met	GAA Glu 650	GAT Asp	CCA Pro	GAC Asp	AAA Lys	GAT Asp 655	GTT Val	196
AGA Arg	GTG Val	GCT Ala	TTT Phe 660	AGT Ser	GGA Gly	AAT Asn	ATC Ile	AAG Lys 665	CAC His	ATA Ile	TTG Leu	GAA Glu	TCC Ser 670	TTG Leu	GAC Asp	2016
TCT Ser	GAA Glu	GAT Asp 675	GGA Gly	TTT Phe	ATA Ile	AAG Lys	GAG Glu 680	CTT Leu	TTT Phe	GTC Val	TTA Leu	AGA Arg 685	ATG Met	AAG Lys	GAA Glu	2064
GCA Ala	TAT Tyr 690	ACA Thr	CAT His	GCC Ala	CAA Gln	ATA Ile 695	TCA Ser	AGA Arg	AAT Asn	AAT Asn	GAG Glu 700	CTG Leu	AAG Lys	GAT Asp	ACC Thr	2112

						GAT Asp										2160
						CAC His										2208
						GCA Ala										2256
GCT Ala	AAA Lys	AGT Ser 755	GTT Val	AAA Lys	CTG Leu	CAA Gln	AGT Ser 760	TTT Phe	TTC Phe	AGC Ser	CAG Gln	TAT Tyr 765	AAG Lys	AAA Lys	CCC Pro	2304
						GAA Glu 775										2352
CTT Leu 785	CCG Pro	AAT Asn	ACT Thr	CCA Pro	TGC Cys 790	CAG Gln	AAT Asn	GCT Ala	GAC Asp	GTG Val 795	CGA Arg	AAA Lys	CAA Gln	GAT Asp	GTG Val 800	2400
						GCT Ala										2448
						CTT Leu										2496
						GCT Ala										2544
						AAA Lys 855										2592
						TAT Tyr										2640
				Leu	Glu	CGT Arg	Ala	Leu	His	Tyr	Leu		Asn			2688
						CTG Leu										2736
AAT Asn	GAA Glu	TTA Leu 915	TTG Leu	CTG Leu	CGT Arg	ATT Ile	GGA Gly 920	GAA Glu	CAC His	TAT Tyr	CAA Gln	CAG Gln 925	GTT Val	TTT Phe	AAT Asn	2784
						TCA Ser 935										2832
						TCA Ser										2880
						TTG Leu										2928

																	•
						GAA Glu											2976
1	TCT Ser	TTG Leu	ATG Met 995	AAG Lys	TTA Leu	ATG Met	GGA Gly	CCC Pro 100	Lys	CAT His	GTC Val	AGT Ser	TCT Ser 100	Val	AGG Arg	GTG Val	3024
			Met			CTG Leu		Thr					Lys				3072
		Glu				AGA Arg 1030	Ala					Val					3120
						TCC Ser					Val					Leu	3168
	CCT Pro	CTT Leu	ATA Ile	CAC His 1060	Ile	CAG Gln	CCT Pro	AAA Lys	GAA Glu 1065	Thr	GCA Ala	GCT Ala	ATC Ile	TTC Phe 1070	His	TAC Tyr	3216
				Glu		AGG Arg			Val					His			3264
			Leu			CAT His		Glu					Lys				3312
		Glu				GAG Glu 1110	Thr					Asp					3360
						AAG Lys					Glu					Arg	3408
					Thr	AGC Ser				Thr					Gln		3456
	AAA Lys	CTG Leu	ATA Ile 1155	Lys	TAT Tyr	GCA Ala	ACA Thr	GAC Asp 1160	Ser	GAA Glu	ACA Thr	GTA Val	GAA Glu 1165	Pro	ATT Ile	ATC Ile	3504
	TCA Ser	CAG Gln 1170	Leu	GTG Val	ACA Thr	GTG Val	CTT Leu 1175	Leu	AAA Lys	GGT Gly	TGC Cys	CAA Gln 1180	Asp	GCA Ala	AAC Asn	TCT Ser	3552
	CAA Gln 1185	Ala	CGG Arg	TTG Leu	CTC Leu	TGT Cys 1190	Gly	GAA Glu	TGT Cys	TTA Leu	GGG Gly 1195	Glu	TTG Leu	GGG Gly	GCG Ala	ATA Ile 1200	3600
	GAT Asp	CCA Pro	GGT Gly	CGA Arg	TTA Leu 1205	GAT A sp	TTC Phe	TCA Ser	ACA Thr	ACT Thr 1210	Glu	ACT Thr	CAA Gln	GGA Gly	AAA Lys 1215	Asp	3648
	TTT Phe	ACA Thr	TTT Phe	GTG Val 1220	Thr	GGA Gly	GTA Val	GAA Glu	GAT Asp 1225	Ser	AGC Ser	TTT Phe	GCC Ala	TAT Tyr 1230	Gly	TTA Leu	3696
,	TTG Leu	ATG Met	GAG Glu 1235	Leu	ACA Thr	AGA Arg	Ala	TAC Tyr 1240	Leu	GCG Ala	TAT Tyr	GCT Ala	GAT Asp 1245	Asn	AGC Ser	CGA Arg	3744

GCT CAA GAT Ala Gln Asp 1250	TCA GCT GCC Ser Ala Ala	TAT GCC AT Tyr Ala Il 1255	T CAG GAG e Gln Glu	TTG CTT TCT Leu Leu Ser 1260	ATT TAT Ile Tyr	3792
GAC TGT AGA Asp Cys Arg 1265	GAG ATG GAG Glu Met Glu 127	Thr Asn Gl	C CCA GGT y Pro Gly 1279	His Gln Leu	TGG AGG Trp Arg 1280	3840
AGA TTT CCT Arg Phe Pro	GAG CAT GTT Glu His Val 1285	CGG GAA AT Arg Glu Il	A CTA GAA e Leu Glu 1290	CCT CAT CTA Pro His Leu	AAT ACC Asn Thr 1295	3888
AGA TAC AAG Arg Tyr Lys	AGT TCT CAG Ser Ser Gln 1300	AAG TCA AC Lys Ser Th	r Asp Trp	TCT GGA GTA Ser Gly Val 131	Lys Lys	3936
CCA ATT TAC Pro Ile Tyr 1315	Leu Ser Lys	TTG GGT AG Leu Gly Se 1320	T AAC TTT r Asn Phe	GCA GAA TGG Ala Glu Trp 1325	TCA GCA Ser Ala	3984
TCT TGG GCA Ser Trp Ala 1330	GGT TAT CTT Gly Tyr Leu	ATT ACA AAG Ile Thr Lys 1335	G GTT ÇGA s Val Arg	CAT GAT CTT His Asp Leu 1340	GCC AGT Ala Ser	4032
AAA ATT TTC Lys Ile Phe 1345	ACC TGC TGT Thr Cys Cys 1350	Ser Ile Me	G ATG AAG t Met Lys 1355	His Asp Phe	AAA GTG Lys Val 1360	4080
ACC ATC TAT Thr Ile Tyr	CTT CTT CCA Leu Leu Pro 1365	CAT ATT CTO	G GTG TAT u Val Tyr 1370	GTC TTA CTG Val Leu Leu	GGT TGT Gly Cys 1375	4128
AAT CAA GAA Asn Gln Glu	GAT CAG CAG Asp Gln Gln 1380	GAG GTT TAT Glu Val Ty: 138	r Ala Glu	ATT ATG GCA Ile Met Ala 1390	Val Leu	4176
AAG CAT GAC Lys His Asp 1395	Asp Gln His	ACC ATA AAT Thr Ile Ass 1400	F ACC CAA n Thr Gln	GAC ATT GCA Asp Ile Ala 1405	TCT GAT Ser Asp	4224
CTG TGT CAA Leu Cys Gln 1410	CTC AGT ACA Leu Ser Thr	CAG ACT GTO Gln Thr Val 1415	G TTC TCC	ATG CTT GAC Met Leu Asp 1420	CAT CTC His Leu	4272
ACA CAG TGG Thr Gln Trp 1425	GCA AGG CAC Ala Arg His 1430	Lys Phe Gl	G GCA CTG n Ala Leu 1435	Lys Ala Glu	AAA TGT Lys Cys 1440	4320
CCA CAC AGC Pro His Ser	AAA TCA AAC Lys Ser Asn 1445	AGA AAT AAG Arg Asn Lys	G GTA GAC s Val Asp 1450	TCA ATG GTA Ser Met Val	TCT ACT Ser Thr 1455	4368
GTG GAT TAT Val Asp Tyr	GAA GAC TAT Glu Asp Tyr 1460	CAG AGT GTA Gln Ser Val 146	l Thr Arg	TTT CTA GAC Phe Leu Asp 1470	Leu Ile	4416
CCC CAG GAT Pro Gln Asp 1475	Thr Leu Ala	GTA GCT TCC Val Ala Ser 1480	TTT CGC r Phe Arg	TCC AAA GCA Ser Lys Ala 1485	TAC ACA Tyr Thr	4464
CGA GCT GTA Arg Ala Val 1490	ATG CAC TTT Met His Phe	GAA TCA TTT Glu Ser Phe 1495	T ATT ACA	GAA AAG AAG Glu Lys Lys 1500	CAA AAT Gln Asn	4512
ATT CAG GAA Ile Gln Glu 1505	CAT CTT GGA His Leu Gly 1510	Phe Leu Glr	G AAA TTG 1 Lys Leu 1515	Tyr Ala Ala	ATG CAT Met His 1520	4560

-119-

			AGA AAG GCA GAA Arg Lys Ala Glu 153	Pro
	Gln Ile Leu Gl		CTT GGC TTG CTG Leu Gly Leu Leu 1550	
	a Cys Tyr Asp Ai		CTA GAA CCA GAC Leu Glu Pro Asp 1565	
			TTA GGT CTT GGT Leu Gly Leu Gly 1580	
			CAT GCT AAC AGG His Ala Asn Arg	
			GAA GCA GCT TGG Glu Ala Ala Trp 161	Lys
	Asp Leu Val Gl		GCA GCA GAT GGA Ala Ala Asp Gly 1630	
	Ser Val Arg Le		TTA TTA TCA GCC Leu Leu Ser Ala 1645	
			AAA CTA GTG AGA Lys Leu Val Arg 1660	
			GAA AGA GGC TCC Glu Arg Gly Ser	
			ATG TTA TGT GAG Met Leu Cys Glu 1699	Leu
	Lys Pro Leu Ph		CCA GGT GAC AGT Pro Gly Asp Ser 1710	
CAA GAA GAT TCT Gln Glu Asp Ser 1715	Leu Asn Trp Va	A GCT CGA CTA 1 Ala Arg Leu 20	GAA ATG ACC CAG Glu Met Thr Gln 1725	AAT 5184 Asn
			CGG AGG GCT TTA Arg Arg Ala Leu 1740	
			GTT GGA GAA TGC Val Gly Glu Cys	
			CAC CAC CAG ACA His His Gln Thr 1779	Ala
TAC AAT GCT CTC Tyr Asn Ala Leu 178	Leu Asn Ala Gl	G GAA TCA CGA y Glu Ser Arg 1785	CTC GCT GAA CTG Leu Ala Glu Leu 1790	TAC 5376 Tyr

-120-

GT0 Val	GAA Glu	AGC Arg 179	, wro	A AAC A Lys	TGG Tr	G CTC D Let	TGC Trp 180	Se:	C AA r Ly	G GG s Gl	T GA y As	T GT p Va 18	l Hi	C CA s Gl	G GCA n Ala		5424
CTA Leu	ATT Ile 181	vai	CTI Leu	CAP Gln	AA Lys	A GGT s Gly 181	' vai	GA Glu	A TT.	A TG u Cy	s Ph	T CC e Pr 20	T GA	A AA' u As:	T GAA n Glu		5472
ACC Thr 182	110	CCT	GAG	GGT Gly	Lys 183	s Asn	ATG Met	TT! Let	A ATO	C CA' Hi:	s Gl	T CG	A GCT	r ATO	G CTA Leu 1840		5520
		O ₁ y	n. g	184	5	. GIU	GIU	inr	185	i Asi	n Ph	e Glu	ı Ser	185			5568
ATT	ATG Met	AAA Lys	AAA Lys 186	IAT	AAG Lys	GAT Asp	GTG Val	ACC Thr 186	Ala	TG(C CT(G CCA	GAA Glu 187	Trp	G GAG Glu		5616
GAT Asp	GGG Gly	CAT His 187	FIIE	TAC Tyr	CTT Leu	GCC Ala	AAG Lys 188	Tyr	ТАТ Туг	GAC	AAA Lys	TTO Leu 188	Met	CCC Pro	ATG Met		5664
GTC Val	ACA Thr 1890	wab	AAC Asn	AAA Lys	ATG Met	GAA Glu 1899	rys	CAA Gln	GGT Gly	GAT Asp	CTC Lev 190	ı Ile	CGG Arg	TAT	ATA Ile		5712
GTT Val 1909	LiC u	CAT His	TTT Phe	GGC Gly	AGA Arg 191	TCT Ser 0	CTA Leu	CAA Gln	TAT Tyr	GGA Gly 191	Asn	CAG Gln	TTC Phe	ATA Ile	TAT Tyr 1920		5760
CAG Gln	TCA Ser	ATG Met	CCA Pro	CGA Arg 1925	mer	TTA Leu	ACT Thr	CTA Leu	TGG Trp 193	Leu	GAT Asp	TAT Tyr	GGT Gly	ACA Thr 193	Lys	-	5808
GCA Ala	TAT Tyr	GAA Glu	TGG Trp 1940	GIU	AAA Lys	GCT Ala	GGC Gly	CGC Arg 1945	Ser	GAT Asp	CGT Arg	GTA Val	CAA Gln 1950	Met	AGG Arg		5856
AAT Asn	Asp	TTG Leu 19 5 5	GIY	AAA Lys	ATA Ile	AAC Asn	AAG Lys 1960	Val	ATC Ile	ACA Thr	GAG Glu	CAT His	Thr	AAC Asn	TAT Tyr		5904
u	GCT Ala 1970	CCA Pro	TAT Tyr	CAA G1n	TTT Phe	TTG Leu 1975	Thr	GCT Ala	TTT Phe	TCA Ser	CAA Gln 198	Leu	ATC Ile	TCT Ser	CGA Arg		5952
ATT 11e 1985	TGT (Cys 1	CAT His	TCT Ser	nis .	GAT Asp 1990	GIU	GTT Val	TTT Phe	GTT Val	GTG Val 1999	Leu	GAT Asp	GGA Gly	AAT Asn	AAT Asn 2000		6000
AGC (CAA (Gln V	GTA ' Val	PHE.	CTA Leu 2005	GCC Ala	TAT (CCT (Pro (Gin	CAA Gln 2010	Ala	ATG Met	TGG Trp	ATG Met	ATG Met 2015	Thr		6048
GCT (TG T	Jer :	AAG ' Lys : 2020	TCA ' Ser :	TCT Ser	TAT (Pro 1	ATG Met 2025	Arg	GTG Val	AAC Asn	AGA Arg	TGC Cys 2030	Lys	GAA Glu		6096
ATC (ieu. F	AAT A Asn 1 2035	AAA (Lys 1	GCT A	ATT Ile	H15 M	ATG A Met 1 2040	AAA . Lys :	AAA Lys	TCC Ser	TTA Leu	GAG Glu 2045	Lys	TTT Phe	GTT Val	(6144
GGA G Gly A	AT G Asp A	CA A	ACT (Thr 1	CGC (Arg I	-eu	ACA C Thr A 2055	SAT A	AAG (CTT Leu	CTA Leu	GAA Glu 2060	Leu	TGC . Cys	AAT Asn	AAA Lys	. (5192

CCG GTT GAT GGA AC Pro Val Asp Gly Se 2065	T AGT TCC ACA r Ser Ser Thr 2070	TTA AGC ATG Leu Ser Met 2075	Ser Thr His	TTT AAA 6240 Phe Lys 2080
ATG CTT AAA AAG CT Met Leu Lys Lys Le 20				
CCT CTA CAA TCA GT Pro Leu Gln Ser Va 2100				Gly Thr
CAT GCT AAC CAT GC His Ala Asn His Al 2115		Pro Phe Pro		
ATT GCA GGG TTT GA Ile Ala Gly Phe As 2130				
CCA AAG AAG ATT TO Pro Lys Lys Ile Se 2145			Lys Phe Tyr	
ATG TGT AAG CCA AA Met Cys Lys Pro Ly 21				
TTC AAT TCC TTG AT Phe Asn Ser Leu Il 2180				Ser Arg
AGA AGA GAA CTT CA Arg Arg Glu Leu Hi 2195		Tyr Ala Val		
GAA TGT GGG ATT AT Glu Cys Gly Ile Il 2210				
ATT CTG ACC AAA CT Ile Leu Thr Lys Le 2225			Tyr Met Thr	
GAA CTT CGC CAG TG Glu Leu Arg Gln Cy 22				
CTC AAA GTA TTC CG Leu Lys Val Phe Ar 2260				Ile Phe
CAT GAG TGG TTT CT His Glu Trp Phe Le 2275		Pro Asp Pro		
AGT AGA TCA GCT TA Ser Arg Ser Ala Ty 2290				
TAT ATT CTG GGG CT Tyr Ile Leu Gly Le 2305			Asn Ile Leu	
TCT TTG ACT GGT GA Ser Leu Thr Gly Gl 23	Cys Val His			

AAG Lys	GGA Gly	GAA Glu	ACC Thr 2340	Phe	GAA Glu	GTT Val	CCA Pro	GAA Glu 234	Ile	GTG Val	CCA Pro	TTT Phe	CGC Arg 2350	Leu	ACT Thr	705
CAT His	AAT Asn	ATG Met 2359	GTT Val	AAT Asn	GGA Gly	ATG Met	GGT Gly 2360	Pro	ATG Met	GGA Gly	ACA Thr	GAG Glu 2365	Gly	CTT Leu	TTT Phe	7104
CGA Arg	AGA Arg 2370	Ala	TGT Cys	GAA Glu	GTT Val	ACA Thr 2375	Met	AGG Arg	CTG Leu	ATG Met	CGT Arg 2380	Asp	CAG Gln	CGA Arg	GAG Glu	7152
CCT Pro 2385	Leu	ATG Met	AGT Ser	GTC Val	TTA Leu 2390	Lys	ACT Thr	TTT Phe	CTA Leu	CAT His 2395	Asp	CCT Pro	CTT Leu	GTG Val	GAA Glu 2400	7200
TGG Trp	AGT Ser	AAA Lys	CCA Pro	GTG Val 2405	Lys	GGG Gly	CAT His	TCC Ser	AAA Lys 2410	Ala	CCA Pro	CTG Leu	AAT Asn	GAA Glu 2415	Thr	7248
GGA Gly	GAA Glu	GTT Val	GTC Val 2420	Asn	GAA Glu	ÀAG Lys	GCC Ala	AAG Lys 2425	Thr	CAT His	GTT Val	CTT Leu	GAC Asp 2430	Ile	GAG Glu	7296
CAG Gln	CGA Arg	CTA Leu 2435	CAA Gln	GGT Gly	GTA Val	ATC Ile	AAG Lys 2440	Thr	CGA Arg	AAT Asn	AGA Arg	GTG Val 2445	Thr	GGA Gly	CTG Leu	7344
CCG Pro	TTA Leu 2450	Ser	ATT Ile	GAA Glu	GGA Gly	CAT His 2455	Val	CAT His	TAC Tyr	CTT Leu	ATA Ile 2460	Gln	GAA Glu	GCT Ala	ACT Thr	7392
GAT Asp 2465	Glu	AAC Asn	TTA Leu	Leu	TGC Cys 2470	Gln	ATG Met	TAT Tyr	CTT Leu	GGT Gly 2475	\mathtt{Trp}	ACT Thr	CCA Pro	Tyr	ATG Met 2480	7440
TGAA	ATGA	T AA	TATG	TAAA	A GA	TATA	GTTA	ATA	ATCI	'AAA	AGTA	AAAA	AA A	AAAA	AAAA	7500
AA																7502

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2480 amino acids
 (B) TYPE: amino acid

 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser

Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu

Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val

Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu.

Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile 65 70 75 80

-123-

Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu 1.05 Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val 200 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn 310 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr Phe Lys Asp Cys Gln His Lys Ser Lys Lys Pro Ser Val Val Ile 360 Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 375 Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 390 Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser 410 Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420

Pro	Trp	Ile 435	Tyr	Ser	His	Ser	Asp 440	Asp	Gly	Cys	Leu	Lys 445	Leu	Thr	Thr
Phe	Ala 450	Ala	Asn	Leu	Leu	Thr 455	Leu	Ser	Cys	Arg	Ile 460	Ser	Asp	Ser	Tyr
Ser 465	Pro	Gln	Ala	Gln	Ser 470	Arg	Суѕ	Val	Phe	Leu 475	Leu	Thr	Leu	Phe	Pro 480
Arg	Arg	Ile	Phe	Leu 485	Glu	Trp	Arg	Thr	Ala 490	Val	Tyr	Asn	Trp	Ala 495	Leu
Gln	Ser	Ser	His 500	Glu	Val	Ile	Arg	Ala 505	Ser	Cys	Val	Ser	Gly 510	Phe	Phe
Ile	Leu	Leu 515	Gln	Gln	Gln	Asn	Ser 520	Cys	Asn	Arg	Val	Pro 525	Lys	Ile	Leu
Ile	Asp 530	Lys	Val	Lys	Asp	Asp 535	Ser	Asp	Ile	Val	Lys 540	Lys	Glu	Phe	Ala
Ser 545	Ile	Leu	Gly	Gln	Leu 550	Val	Cys	Thr	Leu	His 555	Gly	Met	Phe	Tyr	Leu 560
Thr	Ser	Ser	Leu	Thr 565	Glu	Pro	Phe	Ser	Glu 570	His	Gly	His	Val	Asp 575	Leu
Phe	Cys	Arg	Asn 580	Leu	Lys	Ala	Thr	Ser 585	Gln	His	Glu	Cys	Ser 590	Ser	Ser
Gln	Leu	Lys 595	Ala	Ser	Val	Cys	Lys 600	Pro	Phe	Leu	Phe	Leu 605	Leu	Lys	Lys
Lys	Ile 610	Pro	Ser	Pro	Val	Lys 615	Leu	Ala	Phe	Ile	Asp 620	Asn	Leu	His	His
Leu 625	Cys	Lys	His	Leu	Asp 630	Phe	Arg	Glu	Asp	Glu 635	Thr	qaA	Val	Lys	Ala 640
Val	Leu	Gly	Thr	Leu 645	Leu	Asn	Leu	Met	Glu 650	Asp	Pro	Asp	Lys	Asp 655	Val
Arg	Val	Ala	Phe 660	Ser	Gly	Asn	Ile	Lys 665	His	Ile	Leu	Glu	Ser 670	Leu	Asp
Ser	Glu	Asp 675	Gly	Phe	Ile	Lys	Glu 680	Leu	Phe	Val	Leu	Arg 685	Met	Lys	Glu
Ala	Tyr 690	Thr	His	Ala	Gln	Ile 695	Ser	Arg	Asn	Asn	Glu 700	Leu	Lys	Asp	Thr
Leu 705	Ile	Leu	Thr	Thr	Gly 710	Asp	Ile	Gly	Arg	Ala 715	Ala	Lys	Gly	Asp	Leu 720
Val	Pro	Phe	Ala	Leu 725	Leu	His	Leu	Leu	His 730	Cys	Leu	Leu	Ser	Lys 735	Ser
Ala	Ser	Val	Ser 740	Gly	Ala	Ala	Tyr	Thr 745	Glu	Ile	Arg	Ala	Leu 750	Val	Ala
Ala	Lys	Ser 755	Val	Lys	Leu	Gln	Ser 760	Phe	Phe	Ser	Gln	Tyr 765	Lys	Lys	Pro
Ile	Cys 770	Gln	Phe	Leu	Val	Glu 775	Ser	Leu	His	Ser	Ser 780	Gln	Met	Thr	Ala,

WO 97/18323 PCT/US96/19337 -

-125-

Leu 785	Pro	Asn	Thr	Pro	Cys 790	Gln	Asn	Ala	Asp	Val 795	Arg	Lys	Gln	Asp	Val 800
Ala	His	Gln	Arg	Glu 805	Met	Ala	Leu	Asn	Thr 810	Leu	Ser	Glu	Ile	Ala 815	Asn
Val	Phe	Asp	Phe 820	Pro	Asp	Leu	Asn	Arg 825	Phe	Leu	Thr	Arg	Thr 830	Leu	Gln
Val	Leu	Leu 835	Pro	Asp	Leu	Ala	Ala 840	Lys	Ala	Ser	Pro	Ala 845	Ala	Ser	Ala
Leu	Ile 850	Arg	Thr	Leu	Gly	Lys 855	Gln	Leu	Asn	Val	Asn 860	Arg	Arg	Glu	Ile
Leu 865	Ile	Asn	Asn	Phe	Lys 870	Tyr	Ile	Phe	Ser	His 875	Leu	Val	Cys	Ser	Cys 880
Ser	Lys	Asp	Glu	Leu 885	Glu	Arg	Ala	Leu	His 890	Tyr	Leu	Lys	Asn	Glu 895	Thr
Glu	Ile	Glu	Leu 900	Gly	Ser	Leu	Leu	Arg 905	Gln	Asp	Phe	Gln	Gly 910	Leu	His
Asn	Glu	Leu 915	Leu	Leu	Arg	Ile	Gly 920	Glu	His	Tyr	Gln	Gln 925	Val	Phe	Asn
Gly	Leu 930	Ser	Ile	Leu	Ala	Ser 935	Phe	Ala	Ser	Ser	Asp 940	Asp	Pro	Tyr	Gln
Gly 945	Pro	Arg	Asp	Ile	Ile 950	Ser	Pro	Glu	Leu	Met 955	Ala	Asp	Tyr	Leu	Gln 960
Pro	Lys	Leu	Leu	Gly 965	Ile	Leu	Ala	Phe	Phe 970	Asn	Met	Gln	Leu	Leu 975	Ser
Ser	Ser	Val	Gly 980	Ile	Glu	Asp	Lys	Lys 985	Met	Ala	Leu	Asn	Ser 990	Leu	Met
Ser	Leu	Met 995	Lys	Leu	Met	Gly	Pro 1000	-	His	Val	Ser	Ser 1005		Arg	Val
Lys	Met 1010		Thr	Thr	Leu	Arg 1015		Gly	Leu	Arg	Phe 1020	Lys)	Asp	Asp	Phe
Pro 1025		Leu	Cys	Cys	Arg 1030		Trp	Asp	Cys	Phe 1035		Arg	Cys	Leu	Asp 1040
His	Ala	Cys	Leu	Gly 1045		Leu	Leu	Ser	His 1050		Ile	Val	Ala	Leu 1055	
Pro	Leu	Ile	His 1060		Gln	Pro	Lys	Glu 1065		Ala	Ala	Ile	Phe 1070		Tyr
Leu	Ile	Ile 1075		Asn	Arg	Asp	Ala 1080		Gln	Asp	Phe	Leu 1085		Glu	Ile
Tyr	Phe 1090	Leu	Pro	Asp	His	Pro 1095	Glu	Leu	Lys	Lys	Ile 1100	Lys)	Ala	Val	Leu
Gln 1105	Glu	Tyr	Arg	Lys.	Glu 1110		Ser	Glu	Ser	Thr 1115	qaA	Leu	Gln	Thr	Thr 1120
Leu	Gln	Leu	Ser	Met 1125		Ala	Ile	Gln	His		Asn	Val	Asp	Val	

- Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Gln Glu 1140 1145 1150
- Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile 1155 1160 1165
- Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser 1170 1175 1180
- Gln Ala Arg Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile 1185 1190 1195
- Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp 1205 1210 1215
- Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu 1220 1225 1230
- Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg 1235 1240 1245
- Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr 1250 1255 1260
- Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg 1265 1270 1275 1280
- Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr 1285 1290 1295
- Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys 1300 1305 1310
- Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala 1315 1320 1325
- Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1330 1340
- Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val 1345 1350 1355 1360
- Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1365 1370 1375
- Asn Gln Glu Asp Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 1385 1390
- Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395 1400 1405
- Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 1415 1420
- Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 1430 1435 1440
- Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 1450 1455
- Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 1465 1470
- Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 1480 1485

-127-

- Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 1495 1500
- Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 1510 1515 1520
- Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 1530 1535
- Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 1545 1550
- Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 1560 1565
- Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1570 1575 1580
- Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser 1585 1590 1595 1600
- Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys 1605 1610 1615
- Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys 1620 1625 1630
- Ser Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Ser Ala Lys 1635 1640 1645
- Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala 1650 1655 1660
- Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr 1665 1670 1675 1680
- Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu 1685 1690 1695
- Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser 1700 1705 1710
- Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn 1715 1720 1725
- Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu 1730 1740
- Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp 1745 1750 1755 1760
- Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala 1765 1770 1775
- Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr 1780 1785 1790
- Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala 1795 1800 1805
- Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu 1810 1815 1820
- Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu 1825 1830 1835 1840

- Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala 1845 1850 1855
- Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu 1860 1865 1870
- Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met 1875 1880 1885
- Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile 1890 1895 1900
- Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr 1905 1910 1915 1920
- Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys 1925 1930 1935
- Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg 1940 1945 1950
- Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr 1955 1960 1965
- Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg 1970 1975 1980
- Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Asp Gly Asn Asn 1985 1990 1995 2000
- Ser Gln Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr 2005 2010 2015
- Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu 2020 2025 2030
- Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val 2035 2040 2045
- Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Glu Leu Cys Asn Lys 2050 2055 2060
- Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys 2065 2070 2075 2080
- Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile 2085 2090 2095
- Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr 2100 2105 2110
- His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr 2125 2120 2125
- Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys 2130 2140
- Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met 2145 2150 2155 2160
- Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu 2165 2170 2175
- Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg 2180 2185 2190

WO 97/18323 PCT/US96/19337

-129-

Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp 2195 2200 2205

Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro 2210 2215 2220

Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys 2225 2230 2235 2240

Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys 2245 2250 2255

Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe 2260 2265 2270

His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser 2275 2280 2285

Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met Ser Met Val Gly 2290 2295 2300

Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn Ile Leu Phe Asp 2305 2310 2315 2320

Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn Cys Leu Phe Asn 2325 2330 2335

Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro Phe Arg Leu Thr 2340 2345 2350

His Asn Met Val Asn Gly Met Gly Pro Met Gly Thr Glu Gly Leu Phe 2355 2360 2365

Arg Arg Ala Cys Glu Val Thr Met Arg Leu Met Arg Asp Gln Arg Glu 2370 2380

Pro Leu Met Ser Val Leu Lys Thr Phe Leu His Asp Pro Leu Val Glu 2385 2390 2395 2400

Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro Leu Asn Glu Thr 2405 2410 2415

Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val Leu Asp Ile Glu 2420 2425 2430

Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg Val Thr Gly Leu 2435 2440 2445

Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile Gln Glu Ala Thr 2450 2455 2460

Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met 2465 2470 2475 2480

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 878 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 - (b) Torobodi: Tillea
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

-130-

ATCCATTGTG	TTGGAAAGGA	ATGATGAATG	TGGGATTATT	GAATGGGTGA	ACAATACTGC	60
TGGCTTGAGA	CCTATTCTGA	CCAAAATATA	TAAAGAAAAG	GGAGTGTATA	TGACAGGAAA	120
GGAGCTTCGC	CAGTGTATGC	TACCAAAGTC	AGCAGCTTTA	TCTGAAAAAC	TCAAAGTATT	180
CCAAGAATTA	CTCCTGCCCA	GGCATCCTCC	TGTTTTTCAT	GAGTGGTTTC	TGAGAACATT	240
CCCTGATCCT	ACATCATGGT	ACAGTAGCAG	ATCTGCATAT	TGCCGCTCTA	CTGCAGTCAT	300
GTCAATGGTT	GGCTACATCC	TGGGGCTTGG	AGACCGTCAT	GGTGAAAACA	TTCTTTTTGA	360
CTCTTTCACT	GGTGAATGTG	TACATGTAGA	TTTCAACTGT	CTITTTAATA	AGGGAGAAAC	420
GTTTGAAGTT	CCGGAAATTG	TACCATTTCG	ACTGACTCAT	AATATGGTTA	ATGGAATGGG	480
TCCTATGGGA	ACAGAGGGTC	TATTTCGAAG	AGCATGTGAA	GTTACACTGA	GACTGATGAG	540
GGATCAGAGA	GAACCTTTAA	TGAGTGTCTT	AAAGACTTTT	CTACACGATC	CTCTAGTGGA	600
GTGGAGTAAA	CCAGTGAAAG	GACACTCCAA	AGCACCACTG	AATGAAACCG	GGGAAGTTGT	660
CAATGAGAAG	GCCAAGACCC	ATGTTCTTGA	CATTGAACAA	CGACTACAAG	GTGTGATCAA	720
AACCCGAAAT	AGAGTAACAG	GGCTGCCATT	ATCTATTGAA	GGACATGTGC	ATTACCTCAT	780
ACAAGAAGCT	ACTGATGAAA	ACTTACTCTG	TCAGATGTAC	CTTGGTTGGA	CCCCATATAT	840
GTAAAATAAA	ATTATTTCAA	AGAAAAAAA	ААААААА			878

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7935 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..7932

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATG Met 1		CAT His							48
GAG Glu		AGT Ser 20							96

															GTA Val	144
															ACC Thr	192
															CCA Pro 80	240
					GTG Val										TGT Cys	288
					TGG Trp										GCA Ala	336
					TTG Leu											384
					TTT Phe											432
CTC Leu 145	ACA Thr	AAA Lys	GAA Glu	TTA Leu	TTA Leu 150	CAA Gln	CTT Leu	TTT Phe	GAA Glu	GAC Asp 155	TTG Leu	GTT Val	TAC Tyr	CTC Leu	CAT His 160	480
					GGT Gly											528
					TTA Leu											576
					AGT Ser											624
TTA Leu	TTA Leu 210	ATG Met	GTT Val	CTT Leu	ACT Thr	CGT Arg 215	ATT Ile	ATT Ile	GCA Ala	ATT Ile	GTG Val 220	TTT Phe	TTT Phe	AGA Arg	AGG Arg	672
CAA Gln 225	GAA Glu	CTC Leu	TTA Leu	CTT Leu	TGG Trp 230	CAG Gln	ATA Ile	GGT Gly	TGT Cys	GTT Val 235	CTG Leu	CTA Leu	GAG Glu	TAT Tyr	GGT Gly 240	720
AGT Ser	CCA Pro	AAA Lys	ATT Ile	AAA Lys 245	TCC Ser	CTA Leu	GCA Ala	ATT Ile	AGC Ser 250	TTT Phe	TTA Leu	ACA Thr	GAA Glu	CTT Leu 255	TTT Phe	768
CAG Gln	CTT Leu	GGA Gly	GGA Gly 260	CTA Leu	CCA Pro	GCA Ala	CAA Gln	CCA Pro 265	GCT Ala	AGC Ser	ACT Thr	TTT Phe	TTC Phe 270	AGC Ser	TCA Ser	816
TTT	TTG Leu	GAA Glu 275	TTA Leu	TTA Leu	AAA Lys	CAC His	CTT Leu 280	GTA Val	GAA Glu	ATG Met	GAT Asp	ACT Thr 285	GAC Asp	CAA Gln	TTG Leu	864
AAA Lys	CTC Leu 290	TAT Tyr	GAA Glu	GAG Glu	CCA Pro	TTA Leu 295	TCA Ser	AAG Lys	CTG Leu	ATA Ile	AAG Lys 300	ACA Thr	CTA Leu	TTT Phe	CCC Pro	912

TTT Phe 305	GAA Glu	GCA Ala	GAA Glu	GCT Ala	TAT Tyr 310	AGA Arg	TAA Asn	ATT Ile	GAA Glu	CCT Pro 315	GTC Val	TAT Tyr	TTA Leu	AAT Asn	ATG Met 320		960
CTG Leu	CTG Leu	GAA Glu	AAA Lys	CTC Leu 325	TGT Cys	GTC Val	ATG Met	TTT Phe	GAA Glu 330	GAC Asp	GGT Gly	GTG Val	CTC Leu	ATG Met 335	CGG Arg		1008
CTT Leu	AAG Lys	TCT Ser	GAT Asp 340	TTG Leu	CTA Leu	AAA Lys	GCA Ala	GCT Ala 345	TTG Leu	TGC Cys	CAT His	TTA Leu	CTG Leu 350	CAG Gln	TAT Tyr		1056
TTC Phe	CTT Leu	AAA Lys 355	TTT Phe	GTG Val	CCA Pro	GCT Ala	GGG Gly 360	TAT Tyr	GAA Glu	TCT Ser	GCT Ala	TTA Leu 365	CAA Gln	GTC Val	AGG Arg		1104
AAG Lys	GTC Val 370	TAT Tyr	GTG Val	AGA Arg	TAA Asn	ATT Ile 375	TGT Cys	AAA Lys	GCT Ala	CTT Leu	TTG Leu 380	GAT Asp	GTG Val	CTT Leu	GGA Gly		1152
ATT Ile 385	GAG Glu	GTA Val	GAT Asp	GCA Ala	GAG Glu 390	TAC Tyr	TTG Leu	TTG Leu	GGC Gly	CCA Pro 395	CTT Leu	TAT Tyr	GCA Ala	GCT Ala	TTG Leu 400		1200
AAA Lys	ATG Met	GAA Glu	AGT Ser	ATG Met 405	GAA Glu	ATC Ile	ATT Ile	GAG Glu	GAG Glu 410	ATT	CAA Gln	TGC Cys	CAA Gln	ACT Thr 415	CAA Gln		1248
CAG Gln	GAA Glu	AAC Asn	CTC Leu 420	AGC Ser	AGT Ser	AAT Asn	AGT Ser	GAT Asp 425	GGA Gly	ATA Ile	TCA Ser	CCC Pro	AAA Lys 430	AGG Arg	CGT Arg	•	1296
CGT	CTC Leu	AGC Ser 435	TCG Ser	TCT Ser	CTA Leu	AAC Asn	CCT Pro 440	TCT Ser	AAA Lys	AGA Arg	GCA Ala	CCA Pro 445	AAA Lys	CAG Gln	ACT Thr		1344
GAG Glu	GAA Glu 450	ATT Ile	AAA Lys	CAT His	GTG Val	GAC Asp 455	ATG Met	AAC Asn	CAA Gln	AAG Lys	AGC Ser 460	ATA Ile	TTA Leu	TGG Trp	AGT Ser		1392
GCA Ala 465	CTG Leu	AAA Lys	CAG Gln	AAA Lys	GCT Ala 470	GAA Glu	TCC Ser	CTT Leu	CAG Gln	ATT Ile 475	TCC Ser	CTT Leu	GAA Glu	TAC Tyr	AGT Ser 480		1440
GGC Gly	CTA Leu	AAG Lys	AAT Asn	CCT Pro 485	GTT Val	ATT Ile	GAG Glu	ATG Met	TTA Leu 490	GAA Glu	GGA Gly	ATT Ile	GCT Ala	GTT Val 495	GTC Val		1488
TTA Leu	CAA Gln	CTG Leu	ACT Thr 500	GCT Ala	CTG Leu	TGT Cys	ACT Thr	GTT Val 505	CAT	TGT Cys	TCT Ser	CAT His	CAA Gln 510	AAC Asn	ATG Met		1536
AAC Asn	TGC Cys	CGT Arg 515	ACT Thr	TTC Phe	AAG Lys	GAC Asp	TGT Cys 520	CAA Gln	CAT His	AAA Lys	TCC Ser	AAG Lys 525	AAG Lys	AAA Lys	CCT Pro	:	1584
TCT Ser	GTA Val 530	GTG Val	ATA Ile	ACT Thr	TGG Trp	ATG Met 535	TCA Ser	TTG Leu	GAT Asp	TTT Phe	TAC Tyr 540	ACA Thr	AAA Lys	GTG Val	CTT Leu	;	1632
AAG Lys 545	AGC Ser	TGT Cys	AGA Arg	AGT Ser	TTG Leu 550	TTA Leu	GAA Glu	TCT Ser	GTT Val	CAG Gln 555	AAA Lys	CTG Leu	GAC Asp	CTG Leu	GAG Glu 560	•	1680
GCA Ala	ACC Thr	ATT Ile	GAT Asp	AAG Lys 565	GTG Val	GTG Val	AAA Lys	ATT Ile	TAT Tyr 570	GAT Asp	GCT Ala	TTG Leu	ATT Ile	TAT Tyr 575	ATG Met	:	1728

															GGT Gly	1776
					TGG Trp											1824
					GCC Ala											1872
TCA Ser 625	GAT Asp	AGC Ser	TAT Tyr	TCA Ser	CCA Pro 630	CAG Gln	GCA Ala	CAA Gln	TCA Ser	CGA Arg 635	TGT Cys	GTG Val	TTT Phe	CTT Leu	CTG Leu 640	1920
					AGA Arg											1968
AAC Asn	TGG Trp	GCC Ala	CTG Leu 660	CAG Gln	AGC Ser	TCC Ser	CAT His	GAA Glu 665	GTA Val	ATC	CGG Arg	GCT Ala	AGT Ser 670	TGT Cys	GTT Val	2016
					TTA Leu											2064
					GAT Asp											2112
					ATA Ile 710											2160
ATG Met	TTT Phe	TAT Tyr	CTG Leu	ACA Thr 725	AGT Ser	TCT Ser	TTA Leu	ACA Thr	GAA Glu 730	CCT Pro	TTC Phe	TCT Ser	GAA Glu	CAC His 735	GGA Gly	2208
					TGT Cys											2256
TGT Cys	TCA Ser	TCT Ser 755	TCT Ser	CAA Gln	CTA Leu	AAA Lys	GCT Ala 760	TCT Ser	GTC Val	TGC Cys	AAG Lys	CCA Pro 765	TTC Phe	CTT Leu	TTC Phe	2304
CTA Leu	CTG Leu 770	AAA Lys	AAA Lys	AAA Lys	ATA Ile	CCT Pro 775	AGT Ser	CCA Pro	GTA Val	AAA Lys	CTT Leu 780	GCT Ala	TTC Phe	ATA Ile	GAT Asp	2352
AAT Asn 785	CTA Leu	CAT His	CAT His	CTT Leu	TGT Cys 790	AAG Lys	CAT His	CTT Leu	GAT Asp	TTT Phe 795	AGA Arg	GAA Glu	GAT Asp	GAA Glu	ACA Thr 800	2400
GAT Asp	GTA Val	AAA Lys	GCA Ala	GTT Val 805	CTT Leu	GGA Gly	ACT Thr	TTA Leu	TTA Leu 810	AAT Asn	TTA Leu	ATG Met	GAA Glu	GAT Asp 815	CCA Pro	2448
GAC Asp	AAA Lys	GAT Asp	GTT Val 820	AGA Arg	GTG Val	GCT Ala	TTT Phe	AGT Ser 825	GGA Gly	AAT Asn	ATC Ile	AAG Lys	CAC His 830	ATA Ile	TTG Leu	2496
GAA Glu	TCC Ser	TTG Leu 835	GAC Asp	TCT Ser	GAA Glu	GAT Asp	GGA Gly 840	TTT Phe	ATA Ile	AAG Lys	GAG Glu	CTT Leu 845	TTT Phe	GTC Val	TTA Leu	2544

														AAT Asn		2592
CTG Leu 865	AAG Lys	GAT Asp	ACC Thr	TTG Leu	ATT Ile 870	CTT Leu	ACA Thr	ACA Thr	GGG Gly	GAT Asp 875	ATT Ile	GGA Gly	AGG Arg	GCC Ala	GCA Ala 880	2640
														TGT Cys 895		2688
														ATT Ile		2736
														AGC Ser		2784
														TCT Ser		2832
														GTG Val		2880
														TTG Leu 975		2928
														CTT Leu		2976
								Asp					Ala	AGC Ser		3024
		Ser					Thr					Leu		GTC Val		3072
	Arg					Asn					Ile			CAT His		3120
GTC Val	TGT Cys	TCT Ser	TGT Cys	TCC Ser 1045	Lys	GAT Asp	GAA Glu	TTA Leu	GAA Glu 1050	Arg	GCC Ala	CTT Leu	CAT His	TAT Tyr 1055	Leu	3168
AAG Lys	AAT Asn	GAA Glu	ACA Thr 1060	Glu	ATT Ile	GAA Glu	CTG Leu	GGG Gly 1065	Ser	CTG Leu	TTG Leu	AGA Arg	CAA Gln 1070	GAT Asp)	TTC Phe	3216
CAA Gln	GGA Gly	TTG Leu 1075	His	AAT Asn	GAA Glu	TTA Leu	TTG Leu 1080	Leu	CGT Arg	ATT Ile	GGA Gly	GAA Glu 1085	His	TAT Tyr	CAA Gln	3264
CAG Gln	GTT Val 109(Phe	AAT Asn	GGT Gly	TTG Leu	TCA Ser 1095	Ile	CTT Leu	GCC Ala	TCA Ser	TTT Phe 1100	Ala	TCC Ser	AGT Ser	GAT Asp	3312
	Pro					Arg					Pro			ATG Met		3360

			CAA Gln		Lys					Leu						3408
			AGC Ser 114	Ser					Glu					Ala		3456
			ATG Met 5					Leu					His			3504
		Arg	GTG Val				Thr					Gly				3552
	Asp		TTT Phe			Leu					Trp					3600
			GAT Asp		Ala					Leu					Ile	3648
			TTA Leu 1220	Pro					Gln					Ala		3696
ATC Ile	TTC Phe	CAC His 123	TAC Tyr	CTC Leu	ATA Ile	ATT Ile	GAA Glu 1240	Asn	AGG Arg	GAT Asp	GCT Ala	GTG Val 1245	Gln	GAT Asp	TTT Phe	3744
		Glu	ATA Ile				Pro					Leu				3792
	Ala		CTC Leu			Tyr					Ser					3840
			ACT Thr		Gln					Ala					Asn	3888
GTC Val	GAT Asp	GTT Val	CGT Arg 1300	Ile	CAT His	GCT Ala	CTT Leu	ACA Thr 1305	Ser	TTG Leu	AAG Lys	GAA Glu	ACC Thr 1310	Leu	TAT Tyr	3936
AAA Lys	AAT Asn	CAG Gln 1315	GAA Glu	AAA Lys	CTG Leu	ATA Ile	AAG Lys 1320	Tyr	GCA Ala	ACA Thr	GAC Asp	AGT Ser 1325	Glu	ACA Thr	GTA Val	3984
GAA Glu	CCT Pro 1330	Ile	ATC Ile	TCA Ser	CAG Gln	TTG Leu 1335	Val	ACA Thr	GTG Val	CTT Leu	TTG Leu 1340	Lys	GGT Gly	TGC Cys	CAA Gln	4032
GAT Asp 1345	Ala	AAC Asn	TCT Ser	CAA Gln	GCT Ala 1350	Arg	TTG Leu	CTC Leu	TGT Cys	GGG Gly 1355	Glu	TGT Cys	TTA Leu	GGG Gly	GAA Glu 1360	4080
TTG Leu	GGG Gly	GCG Ala	ATA Ile	GAT Asp 1365	Pro	GGT Gly	CGA Arg	TTA Leu	GAT Asp 1370	Phe	TCA Ser	ACA Thr	ACT Thr	GAA Glu 1375	Thr	4128
CAA Gln	GGA Gly	AAA Lys	GAT Asp 1380	Phe '	ACA Thr	TTT Phe	Val	ACT Thr 1385	Gly	GTA Val	GAA Glu	GAT Asp	TCA Ser 1390	Ser	TTT Phe	4176

GCC Ala	TAT Tyr	GGA Gly 139		TTG Leu	ATG Met	GAG Glu	CTA Leu 140	Thr	AGA Arg	GCT Ala	TAC Tyr	CTT Leu 140	Ala	TAT Tyr	GCT Ala	4224
GAT Asp	AAT Asn 141	Ser	CGA Arg	GCT Ala	CAA Gln	GAT Asp 1415	Ser	GCT Ala	GCC Ala	TAT Tyr	GCC Ala 142	Ile	CAG Gln	GAG Glu	TTG Leu	4272
CTT Leu 1425	Ser	ATT Ile	TAT Tyr	GAC Asp	TGT Cys 143	Arg	GAG Glu	ATG Met	GAG Glu	ACC Thr 143	Asn	GGC Gly	CCA Pro	GGT Gly	CAC His 1440	4320
CAA Gln	TTG Leu	TGG Trp	AGG Arg	AGA Arg 144	Phe	CCT Pro	GAG Glu	CAT His	GTT Val 1450	Arg	GAA Glu	ATA Ile	CTA Leu	GAA Glu 145	Pro	4368
CAT His	CTA Leu	AAT Asn	ACC Thr 146	Arg	TAC Tyr	AAG Lys	AGT Ser	TCT Ser 1465	Gln	AAG Lys	TCA Ser	ACC Thr	GAT Asp 147	Trp	TCT Ser	4416
GGA Gly	GTA Val	AAG Lys 147	AAG Lys	CCA Pro	ATT Ile	TAC Tyr	TTA Leu 1480	Ser	AAA Lys	TTG Leu	GGT Gly	AGT Ser 1489	Asn	TTT Phe	GCA Ala	4464
GAA Glu	TGG Trp 149(Ser	GCA Ala	TCT Ser	TGG Trp	GCA Ala 1495	Gly	TAT Tyr	CTT Leu	ATT Ile	ACA Thr 1500	Lys	GTT Val	CGA Arg	CAT His	4512
	Leu		AGT Ser			Phe					Ile					4560
GAT Asp	TTC Phe	AAA Lys	GTG Val	ACC Thr 1525	Ile	TAT Tyr	CTT Leu	CTT Leu	CCA Pro 1530	His	ATT Ile	CTG Leu	GTG Val	TAT Tyr 1535	Val	4608
TTA Leu	CTG Leu	GGT Gly	TGT Cys 154(Asn	CAA Gln	GAA Glu	GAT Asp	CAG Gln 1545	Gln	GAG Glu	GTT Val	TAT	GCA Ala 1550	Glu	ATT Ile	4656
ATG Met	GCA Ala	GTT Val 1555	CTA Leu	AAG Lys	CAT His	GAC Asp	GAT Asp 1560	Gln	CAT His	ACC Thr	ATA Ile	AAT Asn 1565	Thr	CAA Gln	GAC Asp	4704
ATT	GCA Ala 1570	Ser	GAT Asp	C T G Leu	TGT Cys	CAA Gln 1575	Leu	AGT Ser	ACA Thr	CAG Gln	ACT Thr 1580	Val	TTC Phe	TCC Ser	ATG Met	4752
CTT Leu 1585	Asp	CAT His	CTC Leu	ACA Thr	CAG Gln 1590	\mathtt{Trp}	GCA Ala	AGG Arg	CAC His	AAA Lys 1595	Phe	CAG Gln	GCA Ala	CTG Leu	AAA Lys 1600	4800
GCT Ala	GAG Glu	AAA Lys	TGT Cys	CCA Pro 1609	His	AGC Ser	AAA Lys	TCA Ser	AAC Asn 1610	Arg	AAT Asn	AAG Lys	GTA Val	GAC Asp 1615	Ser	4848
ATG Met	GTA Val	TCT Ser	ACT Thr 1620	Val	GAT Asp	TAT Tyr	GAA Glu	GAC Asp 1625	Tyr	CAG Gln	AGT Ser	GTA Val	ACC Thr 1630	Arg	TTT Phe	4896
CTA Leu	GAC Asp	CTC Leu 1635	ATA Ile	CCC Pro	CAG Gln	GAT Asp	ACT Thr 1640	Leu	GCA Ala	GTA Val	GCT Ala	TCC Ser 1645	Phe	CGC Arg	TCC Ser	4944
AAA Lys	GCA Ala 1650	Tyr	ACA Thr	CGA Arg	GCT Ala	GTA Val 1655	Met	CAC His	TTT Phe	GAA Glu	TCA Ser 1660	Phe	ATT Ile	ACA Thr	GAA Glu	4992

AAG AAG CAA AAT Lys Lys Gln Asn 1665	ATT CAG GAA Ile Gln Glu 1670	CAT CTT GGA His Leu Gly	TTT TTA CAG AA Phe Leu Gln Ly: 1675	A TTG TAT 50 s Leu Tyr 1680	040
GCT GCT ATG CAT Ala Ala Met His			Gly Val Ser Ala		880
AAG GCA GAA CCA Lys Ala Glu Pro 1700	Ser Leu Lys	GAA CAG ATC Glu Gln Ile 1705	CTT GAA CAT GA Leu Glu His Gl	Ser Leu	136
GGC TTG CTG AGG Gly Leu Leu Arg 1715					184
GAA CCA GAC CAG Glu Pro Asp Gln 1730		Tyr His Gly			232
GGT CTT GGT CAG Gly Leu Gly Gln 1745					280
GCT AAC AGG TCC Ala Asn Arg Ser			Asn Thr Tyr Arg		328
GCA GCT TGG AAA Ala Ala Trp Lys 1780	Leu Ser Gln			r Leu Ala	376
GCA GAT GGA AAA Ala Asp Gly Lys 1795					424
TTA TCA GCC AAA Leu Ser Ala Lys 1810		Ile Thr Ala			472
CTA GTG AGA GCA Leu Val Arg Ala 1825					520
AGA GGC TCC TAC Arg Gly Ser Tyr			Ile Val Arg Let		568
TTA TGT GAG TTG Leu Cys Glu Leu 1860	Glu His Ser			Ser Pro	616
GGT GAC AGT TCT Gly Asp Ser Ser 1875					664
ATG ACC CAG AAT Met Thr Gln Asn 1890		Ala Lys Glu			712
AGG GCT TTA CTA Arg Ala Leu Leu 1905					760
GGA GAA TGC TGG Gly Glu Cys Trp			Ala Arg Lys Ala		808

CAC CAG ACA GCC TA His Gln Thr Ala Ty 1940	C AAT GCT CTC r Asn Ala Leu	CTT AAT GCA GGG Leu Asn Ala Gly 1945	G GAA TCA CGA C / Glu Ser Arg L 1950	TC 5856 eu
GCT GAA CTG TAC GTG Ala Glu Leu Tyr Va 1955	G GAA AGG GCA l Glu Arg Ala 1960	Lys Trp Leu Trp	G TCC AAG GGT G Ser Lys Gly A 1965	AT 5904 sp
GTT CAC CAG GCA CTA Val His Gln Ala Let 1970	A ATT GTT CTT 1 Ile Val Leu 1975	CAA AAA GGT GTT Gln Lys Gly Val 198	. Glu Leu Cys Pi	TT 5952 he
CCT GAA AAT GAA ACC Pro Glu Asn Glu Thi 1985	C CCA CCT GAG Pro Pro Glu 1990	GGT AAG AAC ATG Gly Lys Asn Met 1995	Leu Ile His G	GT 6000 ly 000
CGA GCT ATG CTA CTA Arg Ala Met Leu Leu 200	i val Gly Arg	TTT ATG GAA GAA Phe Met Glu Glu 2010	ACA GCT AAC TT Thr Ala Asn Ph 2015	FT 6048 ne
GAA AGC AAT GCA ATT Glu Ser Asn Ala Ile 2020	e Met Lys Lys	TAT AAG GAT GTG Tyr Lys Asp Val 2025	ACC GCG TGC CT Thr Ala Cys Le 2030	rg 6096 eu
CCA GAA TGG GAG GAT Pro Glu Trp Glu Asp 2035	GGG CAT TTT CONTROL GIVE HIS Phe CONTROL GIVE HIS Phe CONTROL GIVE TO THE CONTROL GIVE THE CONTROL GIVE TO THE CONTROL GIVE THE CONTROL GIVE TO THE CONTROL GIVE THE CONTROL G	TAC CTT GCC AAG Tyr Leu Ala Lys	TAC TAT GAC AA Tyr Tyr Asp Ly 2045	AA 6144 vs
TTG ATG CCC ATG GTC Leu Met Pro Met Val 2050	ACA GAC AAC A Thr Asp Asn I 2055	AAA ATG GAA AAG Lys Met Glu Lys 206	Gln Gly Asp Le	CC 6192
ATC CGG TAT ATA GTT Ile Arg Tyr Ile Val 2065	CTT CAT TTT C Leu His Phe C 2070	GGC AGA TCT CTA Gly Arg Ser Leu 2075	Gln Tyr Gly As	AT 6240 en 80
CAG TTC ATA TAT CAG Gln Phe Ile Tyr Gln 208	Ser Met Pro A	CGA ATG TTA ACT Arg Met Leu Thr 2090	CTA TGG CTT GA Leu Trp Leu As 2095	T 6288 P
TAT GGT ACA AAG GCA Tyr Gly Thr Lys Ala 2100	Tyr Glu Trp G	GAA AAA GCT GGC Glu Lys Ala Gly 2105	CGC TCC GAT CG Arg Ser Asp Ar 2110	T 6336 g
GTA CAA ATG AGG AAT Val Gln Met Arg Asn 2115	GAT TTG GGT A Asp Leu Gly I 2120	AAA ATA AAC AAG Lys Ile Asn Lys	GTT ATC ACA GA Val Ile Thr Gl 2125	G 6384 u
CAT ACA AAC TAT TTA His Thr Asn Tyr Leu 2130	GCT CCA TAT C Ala Pro Tyr G 2135	CAA TTT TTG ACT In Phe Leu Thr 2140	Ala Phe Ser Gla	A 6432 n
TTG ATC TCT CGA ATT Leu Ile Ser Arg Ile 2145	TGT CAT TCT C Cys His Ser H 2150	CAC GAT GAA GTT lis Asp Glu Val 2155	TTT GTT GTC TTC Phe Val Val Let 210	u
ATG GAA ATA ATA GCC Met Glu Ile Ile Ala . 2169	Lys Val Phe L	TA GCC TAT CCT eu Ala Tyr Pro 2170	CAA CAA GCA ATO Gln Gln Ala Med 2175	G 6528 t
TGG ATG ATG ACA GCT Trp Met Met Thr Ala 2180	Val Ser Lys S	CA TCT TAT CCC er Ser Tyr Pro 185	ATG CGT GTG AAG Met Arg Val Ass 2190	C 6576 n
AGA TGC AAG GAA ATC Arg Cys Lys Glu Ile 2195	CTC AAT AAA G Leu Asn Lys A 2200	CT ATT CAT ATG la Ile His Met	AAA AAA TCC TTA Lys Lys Ser Let 2205	A 6624

GAG AAG TTT GTT GG Glu Lys Phe Val Gl 2210	A GAT GCA ACT / Asp Ala Thr 2215	CGC CTA ACA Arg Leu Thr	GAT AAG CTT CTA Asp Lys Leu Leu 2220	GAA 6672 Glu
TTG TGC AAT AAA CC Leu Cys Asn Lys Pr 2225			Thr Leu Ser Met	
ACT CAT TTT AAA AT Thr His Phe Lys Me 22	Leu Lys Lys	CTG GTA GAA Leu Val Glu 2250	GAA GCA ACA TTT Glu Ala Thr Phe 225	Ser
GAA ATC CTC ATT CC Glu Ile Leu Ile Pr 2260	r CTA CAA TCA o Leu Gln Ser	GTC ATG ATA Val Met Ile 2265	CCT ACA CTT CCA Pro Thr Leu Pro 2270	TCA 6816 Ser
ATT CTG GGT ACC CA Ile Leu Gly Thr Hi 2275	r GCT AAC CAT s Ala Asn His 2280	Ala Ser His	GAA CCA TTT CCT Glu Pro Phe Pro 2285	GGA 6864 Gly
CAT TGG GCC TAT AT His Trp Ala Tyr Il 2290	GCA GGG TTT Ala Gly Phe 2295	GAT GAT ATG Asp Asp Met	GTG GAA ATT CTT Val Glu Ile Leu 2300	GCT 6912 Ala
TCT CTT CAG AAA CC Ser Leu Gln Lys Pr 2305	A AAG AAG ATT D Lys Lys Ile 2310	TCT TTA AAA Ser Leu Lys 231	Gly Ser Asp Gly	AAG 6960 Lys 2320
TTC TAC ATC ATG AT Phe Tyr Ile Met Me 23	Cys Lys Pro			Cys
AGA CTA ATG GAA TT Arg Leu Met Glu Ph 2340				
GCA GAG TCT CGT AG Ala Glu Ser Arg Ar 2355	A AGA GAA CTT g Arg Glu Leu 2360	His Ile Arg	ACA TAT GCA GTT Thr Tyr Ala Val 2365	ATT 7104 Ile
CCA CTA AAT GAT GA Pro Leu Asn Asp Gl 2370	A TGT GGG ATT L Cys Gly Ile 2375	ATT GAA TGG Ile Glu Trp	GTG AAC AAC ACT Val Asn Asn Thr 2380	GCT 7152 Ala
GGT TTG AGA CCT AT Gly Leu Arg Pro Il 2385			Glu Lys Gly Val	
ATG ACA GGA AAA GA Met Thr Gly Lys Gl 24	ı Leu Arg Gln			Ala
TTA TCT GAA AAA CT Leu Ser Glu Lys Le 2420				
CCT CCT ATT TTT CA Pro Pro Ile Phe Hi 2435		Leu Arg Thr		
TCA TGG TAC AGT AG Ser Trp Tyr Ser Se 2450				
TCA ATG GTT GGT TA Ser Met Val Gly Ty 2465			Arg His Gly Glu	

-140-

ATT Ile	CTC Leu	TTT Phe	GAT Asp	TCT Ser 248	Leu	ACT Thr	GGT Gly	GAA Glu	TGC Cys 249	GTA Val 0	CAT His	GTA Val	GAT Asp	TTC Phe 249	Asn	7488
TGT Cys	CTT Leu	TTC Phe	AAT Asn 2500	Lys	GGA Gly	GAA Glu	ACC Thr	TTT Phe 250	Glu	GTT Val	CCA Pro	GAA Glu	ATT Ile 251	Val	CCA Pro	7536
TTT Phe	CGC Arg	CTG Leu 2515	Thr	CAT His	AAT Asn	ATG Met	GTT Val 2520	Asn	GGA Gly	ATG Met	GGT Gly	CCT Pro 2529	Met	GGA Gly	ACA Thr	7584
GAG Glu	GGT Gly 2530	Leu	TTT Phe	CGA Arg	AGA Arg	GCA Ala 2535	Cys	GAA Glu	GTT Val	ACA Thr	ATG Met 254(Arg	CTG Leu	ATG Met	CGT Arg	7632
GAT Asp 2545	Gln	CGA Arg	GAG Glu	CCT Pro	TTA Leu 2550	Met	AGT Ser	GTC Val	TTA Leu	AAG Lys 2555	Thr	TTT Phe	CTA Leu	CAT His	GAT Asp 2560	7680
CCT Pro	CTT Leu	GTG Val	GAA Glu	TGG Trp 2565	Ser	AAA Lys	CCA Pro	GTG Val	AAA Lys 2570	GGG Gly	CAT His	TCC Ser	AAA Lys	GCG Ala 2575	Pro	7728
CTG Leu	AAT Asn	GAA Glu	ACT Thr 2580	Gly	GAA Glu	GTT Val	GTC Val	AAT Asn 2585	Glu	AAG Lys	GCC Ala	AAG Lys	ACC Thr 2590	His	GTT Val	7 77 6
CTT Leu	GAC Asp	ATT Ile 2595	Glu	CAG Gln	CGA Arg	CTA Leu	CAA Gln 2600	Gly	GTA Val	ATC	AAG Lys	ACT Thr 2605	Arg	AAT Asn	AGA Arg	7824
GTG Val	ACA Thr 2610	Gly	CTG Leu	CCG Pro	TTA Leu	TCT Ser 2615	Ile	GAA Glu	GGA Gly	CAT His	GTG Val 2620	His	TAC Tyr	CTT Leu	ATA Ile	7872
CAA Gln 2625	Glu	GCT Ala	ACT Thr	GAT Asp	GAA Glu 2630	Asn	TTA Leu	CTA Leu	TGC Cys	CAG Gln 2635	Met	TAT Tyr	CTT Leu	GGT Gly	TGG Trp 2640	7920
	CCA Pro			TGA												7935

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2644 amino acids (B) TYPE: amino acid

 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gly Glu His Gly Leu Glu Leu Ala Ser Met Ile Pro Ala Leu Arg

Glu Leu Gly Ser Ala Thr Pro Glu Glu Tyr Asn Thr Val Val Gln Lys 20

Pro Arg Gln Ile Leu Cys Gln Phe Ile Asp Arg İle Leu Thr Asp Val

Asn	Val 50	Val	Ala	Val	Glu	Leu 55	Val	Lys	Lys	Thr	Asp 60	Ser	Gln	Pro	Thr
Ser 65	Val	Met	Leu	Leu	Asp 70	Phe	Ile	Gln	His	Ile 75	Met	Lys	Ser	Ser	Pro 80
Leu	Met	Phe	Val	Asn 85	Val	Ser	Gly	Ser	His 90	Glu	Arg	Lys	Gly	Ser 95	Сув
Ile	Glu	Phe	Ser 100	Asn	Trp	Ile	Ile	Thr 105	Arg	Leu	Leu	Arg	Ile 110	Ala	Ala
Thr	Pro	Ser 115	Cys	His	Leu	Leu	His 120	Lys	Lys	Ile	Cys	Glu 125	Val	Ile	Cys
Ser	Leu 130	Leu	Phe	Leu	Phe	Lys 135	Ser	Lys	Ser	Pro	Ala 140	Ile	Phe	Gly	Val
Leu 145	Thr	Lys	Glu	Leu	Leu 150	Gln	Leu	Phe	Glu	Asp 155	Leu	Val	Tyr	Leu	His 160
Arg	Arg	Asn	Val	Met 165	Gly	His	Ala	Val	Glu 170	Trp	Pro	Val	Val	Met 175	Ser
Arg	Phe	Leu	Ser 180	Gln	Leu	Asp	Glu	His 185	Met	Gly	Tyr	Leu	Gln 190	Ser	Ala
Pro	Leu	Gln 195	Leu	Met	Ser	Met	Gln 200	Asn	Leu	Glu	Phe	Ile 205	Glu	Val	Thr
Leu	Leu 210	Met	Val	Leu	Thr	Arg 215	Ile	Ile	Ala	Ile	Val 220	Phe	Phe	Arg	Arg
Gln 225	Glu	Leu	Leu	Leu	Trp 230	Gln	Ile	Gly	Cys	Val 235	Leu	Leu	Glu	Tyr	Gly 2 4 0
Ser	Pro	Lys	Ile	Lys 245	Ser	Leu	Ala	Ile	Ser 250	Phe	Leu	Thr	Glu	Leu 255	Phe
Gln	Leu	Gly	Gly 260	Leu	Pro	Ala	Gln	Pro 265	Ala	Ser	Thr	Phe	Phe 270	Ser	Ser
Phe	Leu	Glu 275	Leu	Leu	Lys	His	Leu 280	Val	Glu	Met	Asp	Thr 285	Asp	Gln	Leu
Lys	Leu 290	Tyr	Glu	Glu	Pro	Leu 295	Ser	Lys	Leu	Ile	Lys 300	Thr	Leu	Phe	Pro
Phe 305	Glu	Ala	Glu	Ala	Tyr 310	Arg	Asn	Ile	Glu	Pro 315	Val	Tyr	Leu	Asn	Met 320
Leu	Leu	Glu	Lys	Leu 325	Cys	Val	Met	Phe	Glu 330	Asp	Gly	Val	Leu	Met 335	Arg
Leu	Lys	Ser	Asp 340	Leu	Leu	Lys	Ala	Ala 345	Leu	Cys	His	Leu	Leu 350	Gln	Tyr
Phe	Leu	Lys 355	Phe	Val	Pro	Ala	Gly 360	Tyr	Glu	Ser	Ala	Leu 365	Gln	Val	Arg
Lys	Val 370	Tyr	Val	Arg	Asn	Ile 375	Cys	Lys	Ala	Leu	Leu 380	Asp	Val	Leu	Gly
Ile 385	Glu	Val	Asp	Ala	Glu 390	Tyr	Leu	Leu	Gly	Pro 395	Leu	Tyr	Ala	Ala	Leu 400

Lys	Met	Glu	Ser	Met 405	Glu	Ile	Ile	Glü	Glu 410	Ile	Gln	Cys	Gln	Thr 415	Gln
Gln	Glu	Asn	Leu 420	Ser	Ser	Asn	Ser	Asp 425	Gly	Ile	Ser	Pro	Lys 430	Arg	Arg
Arg	Leu	Ser 435	Ser	Ser	Leu	Asn	Pro 440	Ser	Lys	Arg	Ala	Pro 445	Lys	Gln	Thr
Glu	Glu 450	Ile	Lys	His	Val	Asp 455	Met	Asn	Gln	Lys	Ser 460	Ile	Leu	Trp	Ser
Ala 465	Leu	Lys	Gln	Lys	Ala 470	Glu	Ser	Leu	Gln	Ile 475	Ser	Leu	Glu	Tyr	Ser 480
Gly	Leu	Lys	Asn	Pro 485	Val	Ile	Glu	Met	Leu 490	Glu	Gly	Ile	Ala	Val 495	Val
Leu	Gln	Leu	Thr 500	Ala	Leu	Cys	Thr	Val 505	His	Cys	Ser	His	Gln 510	Asn	Met
Asn	Cys	Arg 515	Thr	Phe	Lys	Asp	Cys 520	Gln	His	Lys	Ser	Lys 525	Lys	Lys	Pro
Ser	Val 530	Val	Ile	Thr	Trp	Met 535	Ser	Leu	Asp	Phe	Tyr 540	Thr	Lys	Val	Leu
Lys 545	Ser	Cys	Arg	Ser	Leu 550	Leu	Glu	Ser	Val	Gln 555	Lys	Leu	Asp	Leu	Glu 560
Ala	Thr	Ile	Asp	Lys 565	Val	Val	Lys	Ile	Tyr 570	Asp	Ala	Leu	Ile	Tyr 575	Met
Gln	Val	naA	Ser 580	Ser	Phe	Glu	Asp	His 585	Ile	Leu	Glu	Asp	Leu 590	Cys	Gly
Met	Leu	Ser 595	Leu	Pro	Trp	Ile	Tyr 600	Ser	His	Ser	qaA	Asp 605	Gly	Cys	Leu
Lys	Leu 610	Thr	Thr	Phe	Ala	Ala 615	Asn	Leu	Leu	Thr	Leu 620	Ser	Cys	Arg	Ile
Ser 625	Asp	Ser	Tyr	Ser	Pro 630	Gln	Ala	Gln	Ser	Arg 635	Cys	Val	Phe	Leu	Leu 640
Thr	Leu	Phe	Pro	Arg 645	Arg	Ile	Phe	Leu	Glu 650	Trp	Arg	Thr	Ala	Val 655	Tyr
Asn	Trp	Ala	Leu 660	Gln	Ser	Ser	His	Glu 665	Val	Ile	Arg	Ala	Ser 670	Cys	Val
Ser	Gly	Phe 675	Phe	Ile	Leu	Leu	Gln 680	Gln	Gln	Asn	Ser	Cys 685	Asn	Arg	Val
Pro	Lys 690	Ile	Leu	Ile	Asp	Lys 695	Val	Lys	Asp	qaA	Ser 700	Asp	Ile	Val	Lys
Lys 705	Glu	Phe	Ala	Ser	Ile 710	Leu	Gly	Gln	Leu	Val 715	Cys	Thr	Leu	His	Gly 720
Met	Phe	Tyr	Leu	Thr 725	Ser	Ser	Leu	Thr	Glu 730	Pro	Phe	Ser	Glu	His 735	Gly
His	Val	Asp	Leu 740	Phe	Cys	Arg	Asn	Leu 745	Lys	Ala	Thr	Ser	Gln 750	His	Glu

Cys	Ser	Ser 755		Gln	Leu	Lys	760	Ser	· Val	Cys	Lys	765		Lei	ı Pho
Leu	Leu 770		Lys	Lys	Ile	Pro 775		Pro	Val	Lys	780		Phe	Ile	a Ası
Asn 785		His	His	Leu	Cys 790		His	Leu	Asp	Phe 795		Glu	Asp	Glu	Th:
Asp	Val	Lys	Ala	Val 805		Gly	Thr	Leu	Leu 810		Leu	Met	Glu	Asp 815	
Asp	Lys	Asp	Val 820		Val	Ala	Phe	Ser 825		Asn	Ile	Lys	His 830		Let
Glu	Ser	Leu 835	Asp	Ser	Glu	Asp	Gly 840	Phe	Ile	Lys	Glu	Leu 845		Val	Let
Arg	Met 850	Lys	Glu	Ala	Tyr	Thr 855	His	Ala	Gln	Ile	Ser 860	Arg	Asn	Asn	Glu
Leu 865	Lys	Asp	Thr	Leu	Ile 870	Leu	Thr	Thr	Gly	Asp 875		Gly	Arg	Ala	Ala 880
Lys	Gly	Asp	Leu	Val 885	Pro	Phe	Ala	Leu	Leu 890	His	Leu	Leu	His	Cys 895	
Leu	Ser	Lys	Ser 900	Ala	Ser	Val	Ser	Gly 905	Ala	Ala	Tyr	Thr	Glu 910	Ile	Arg
Ala	Leu	Val 915	Ala	Ala	Lys	Ser	Val 920	Lys	Leu	Gln	Ser	Phe 925	Phe	Ser	Gln
Tyr	Lys 930	Lys	Pro	Ile	Cys	Gln 935	Phe	Leu	Val	Glu	Ser 940	Leu	His	Ser	Ser
Gln 945	Met	Thr	Ala	Leu	Pro 950	Asn	Thr	Pro	Cys	Gln 955	Asn	Ala	Asp	Val	Arg 960
Lys	Gln	Asp	Val	Ala 965	His	Gln	Arg	Glu	Met 970	Ala	Leu	Asn	Thr	Leu 975	Ser
Glu	Ile	Ala	Asn 980	Val	Phe	Asp	Phe	Pro 985	Asp	Leu	Asn	Arg	Phe 990	Leu	Thr
Arg	Thr	Leu 995	Gln	Val	Leu	Leu	Pro 1000	Asp O	Leu	Ala	Ala	Lys 100		Ser	Pro
Ala	Ala 1010	Ser	Ala	Leu	Ile	Arg 1015	Thr	Leu	Gly	Lys	Gln 1020		Asn	Val	Asn
Arg 1025	Arg	Glu	Ile	Leu	Ile 1030	Asn)	Asn	Phe	Lys	Tyr 1035		Phe	Ser	His	Leu 104
Val	Cys	Ser	Cys	Ser 1049	Lys	Asp	Glu	Leu	Glu 1050	Arg	Ala	Leu	His	Tyr 1059	
Lys	Asn	Glu	Thr 1060	Glu)	Ile	Glu	Leu	Gly 1065		Leu	Leu	Arg	Gln 1070		Phe
Gln	Gly	Leu 1075	His	Asn	Glu	Leu	Leu 1080	Leu)	Arg	Ile	Gly	Glu 1085		Tyr	Gln
Gln	Val 1090	Phe	Asn	Gly	Leu	Ser 1095	Ile	Leu	Ala	Ser	Phe		Ser	Ser	Asp

Asp 110		Tyr	Gln	Gly	Pro 111		Asp	Ile	Ile	Ser 1111		Glu	Leu	Met	Ala 1120
Asp	Tyr	Leu	Gln	Pro 1129	Lys 5	Leu	Leu	Gly	Ile 1130		Ala	Phe	Phe	Asn 1139	
Gln	Leu	Leu	Ser 114(Ser	Val	Gly	Ile 1145		Asp	Lys	Lys	Met 1150		Leu
Asn	Ser	Leu 1155		Ser	Leu	Met	Lys 1160		Met	Gly	Pro	Lys 1165		Val	Ser
Ser	Val 1170		Val	Lys	Met	Met 1175		Thr	Leu	Arg	Thr 1180		Leu	Arg	Phe
Lys 1189		Asp	Phe	Pro	Glu 1190		Сув	Сув	Arg	Ala 1199		Asp	Cys	Phe	Val 1200
Arg	Cys	Leu	qaA	His 1205	Ala	Cys	Leu	Gly	Ser 1210		Leu	Ser	His	Val 1215	
Val	Ala	Leu	Leu 1220		Leu	Ile	His	Ile 1225	Gln	Pro	Lys	Glu	Thr 1230		Ala
Ile	Phe	His 1235		Leu	Ile	Ile	Glu 1240		Arg	Asp	Ala	Val 1245		Asp	Phe
Leu	His 1250		Ile	Tyr	Phe	Leu 1255		Asp	His	Pro	Glu 1260		Lys	Lys	Ile
Lys 1265		Val	Leu	Gln	Glu 1270		Arg	Lys	Glu	Thr 1275		Glu	Ser	Thr	Asp 1280
Leu	Gln	Thr	Thr	Leu 1285	Gln	Leu	Ser	Met	Lys 1290		Ile	Gln	His	Glu 1295	
Val	Asp	Val	Arg 1300		His	Ala	Leu	Thr 1305		Leu	Lys	Glu	Thr 1310		Tyr
Lys	Asn	Gln 1315		Lys	Leu	Ile	Lys 1320		Ala	Thr	Asp	Ser 1325		Thr	Val
Glu	Pro 1330		Ile	Ser	Gln	Leu 1335		Thr	Val	Leu	Leu 1340	_	Gly	Cys	Gln
Asp 1345					Ala 1350								Leu		Glu 1360
Leu	Gly	Ala	Ile	Asp 1365	Pro	Gly	Arg	Leu	Asp 1370		Ser	Thr	Thr	Glu 1375	
Gln	Gly	Lys	Asp 1380		Thr	Phe	Val	Thr 1385		Val	Glu	Asp	Ser 1390		Phe
Ala	Tyr	Gly 1395		Leu	Met	Glu	Leu 1400		Arg	Ala	Tyr	Leu 1405		Tyr	Ala
Asp	Asn 1410		Arg	Ala	Gln	Asp 1415		Ala	Ala	Tyr	Ala 1420		Gln	Glu	Leu
Leu 1425		Ile	Tyr	qaA	Сув 1430		Glu	Met	Glu	Thr 1435		Gly	Pro	Gly	His 1440

Gln Leu Trp Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro 1445 1450 1455

- His Leu Asn Thr Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser 1460 1465 1470
- Gly Val Lys Lys Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala 1475 1480 1485
- Glu Trp Ser Ala Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His 1490 1495 1500
- Asp Leu Ala Ser Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His 1505 1510 1515 1520
- Asp Phe Lys Val Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val 1525 1530 1535
- Leu Leu Gly Cys Asn Gln Glu Asp Gln Glu Val Tyr Ala Glu Ile 1540 1545 1550
- Met Ala Val Leu Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp 1555 1560 1565
- Ile Ala Ser Asp Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met 1570 1575 1580
- Leu Asp His Leu Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys 1585 1590 1595 1600
- Ala Glu Lys Cys Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser 1605 1610 1615
- Met Val Ser Thr Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe 1620 1630
- Leu Asp Leu Ile Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser 1635 1640 1645
- Lys Ala Tyr Thr Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu 1650 1660
- Lys Lys Gln Asn Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr 1665 1670 1675 1680
- Ala Ala Met His Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg 1685 1690 1695
- Lys Ala Glu Pro Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu 1700 1705 1710
- Gly Leu Leu Arg Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu 1715 1720 1725
- Glu Pro Asp Gln Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu 1730 1740
- Gly Leu Gly Gln Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His 1745 1750 1755 1760
- Ala Asn Arg Ser Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu 1765 1770 1775
- Ala Ala Trp Lys Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala 1780 1785 1790
- Ala Asp Gly Lys Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu 1795 1800 1805

- Leu Ser Ala Lys Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys 1810 1815 1820
- Leu Val Arg Ala Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu 1825 1830 1835 1840
- Arg Gly Ser Tyr Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met 1845 1850 1855
- Leu Cys Glu Leu Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro 1860 1865 1870
- Gly Asp Ser Ser Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu 1875 1880 1885
- Met Thr Gln Asn Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg 1890 1895 1900
- Arg Ala Leu Leu Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val 1905 1910 1915 1920
- Gly Glu Cys Trp Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His 1925 1930 1935
- His Gln Thr Ala Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu 1940 1945 1950
- Ala Glu Leu Tyr Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp 1955 1960 1965
- Val His Gln Ala Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe 1970 1975 1980
- Pro Glu Asn Glu Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly 1985 1990 1995 2000
- Arg Ala Met Leu Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe 2005 2010 2015
- Glu Ser Asn Ala Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu 2020 2025 2030
- Pro Glu Trp Glu Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys 2035 2040 2045
- Leu Met Pro Met Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu 2050 2060
- Ile Arg Tyr Ile Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn 2065 2070 2075 2080
- Gln Phe Ile Tyr Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp 2085 2090 2095
- Tyr Gly Thr Lys Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg 2100 2105 2110
- Val Gln Met Arg Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu 2115 2120 2125
- His Thr Asn Tyr Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln 2130 2135 2140
- Leu Ile Ser Arg Ile Cys His Ser His Asp Glu Val Phe Val Val Leu 2145 2150 2155 2160

- Met Glu Ile Ile Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met 2165 2170 2175
- Trp Met Met Thr Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn 2180 2185 2190
- Arg Cys Lys Glu Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu 2195 2200 2205
- Glu Lys Phe Val Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu 2210 2215 2220
- Leu Cys Asn Lys Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser 2225 2230 2235 2240
- Thr His Phe Lys Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser 2245 2250 2255
- Glu Ile Leu Ile Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser 2260 2265 2270
- Ile Leu Gly Thr His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly 2275 2280 2285
- His Trp Ala Tyr Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala 2290 2295 2300
- Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys 2305 2310 2315 2320
- Phe Tyr Ile Met Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys 2325 2330 2335
- Arg Leu Met Glu Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp 2340 2345 2350
- Ala Glu Ser Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile 2355 2360 2365
- Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala 2370 2375 2380
- Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr 2385 2390 2395 2400
- Met Thr Gly Lys Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala 2405 2410 2415
- Leu Ser Glu Lys Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His 2420 2425 2430
- Pro Pro Ile Phe His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr 2435 2440 2445
- Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys Arg Ser Thr Ala Val Met 2450 2460
- Ser Met Val Gly Tyr Ile Leu Gly Leu Gly Asp Arg His Gly Glu Asn 2465 2470 2475 2480
- Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys Val His Val Asp Phe Asn 2485 2490 2495
- Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu Val Pro Glu Ile Val Pro 2500 2505 2510

-148-

Phe	Arg	Leu 251	Thr 5	His	Asn	Met	Val 252	Asn O	Gly	Met	Gly	Pro 2525		Gly	Thr
Glu	Gly 2530	Leu)	Phe	Arg	Arg	Ala 2535	Cys	Glu	Val	Thr	Met 2540		Leu	Met	Arg
Asp	Gln	Ara	Glu	Pro	Leu	Met	Sar	Va 1	Lau	Lic	The	Dho	7	174 ~	3

Met Ser Val Leu Lys Thr Phe Leu His Asp 2555 2560 Asp Gin Arg Giu Pro Leu M 2545 2550

Pro Leu Val Glu Trp Ser Lys Pro Val Lys Gly His Ser Lys Ala Pro 2565

Leu Asn Glu Thr Gly Glu Val Val Asn Glu Lys Ala Lys Thr His Val 2585

Leu Asp Ile Glu Gln Arg Leu Gln Gly Val Ile Lys Thr Arg Asn Arg

Val Thr Gly Leu Pro Leu Ser Ile Glu Gly His Val His Tyr Leu Ile 2615

Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys Gln Met Tyr Leu Gly Trp 2635

Thr Pro Tyr Met

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7624 base pairs

 - (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 333..7562
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTTGTGAAGA GAATGTTTTA CACTCTTGTT AGTGAAGTTT ATTCTTTAAA AGTCAATCGT	60
CAAGGATTTA GCAAATGAAT TAGCACTTCG GATATACTTG TTTATTTAAT ATCTTTTTTG	120
TTTATTTCAA AGAATTCAGT AATTGGATCA TAACGAGACT TCTGCGGATT GCAGCAACTC	180
CCTCCTGTCA TTTGTTACAC AAGAAAATCT GTGAAGTCAT CTGTTCATTA TTATTTCTTT	240
TTAAAAGCAA GAGTCCTGCT ATTTTTGGGG TACTCACAAA AGAATTATTA CAACTTTTTG	300
AAGACTTGGT TTACCTCCAT AGAAGAAATG TG ATG GGT CAT GCT GTG GAA TGG Met Gly His Ala Val Glu Trp 1 5	353
CCA GTG GTC ATG AGC CGA TTT TTA AGT CAA TTA GAT GAA CAC ATG GGA Pro Val Val Met Ser Arg Phe Leu Ser Gln Leu Asp Glu His Met Gly 10 15 20	401
TAT TTA CAA TCA GCT CCT TTG CAG TTG ATG AGT ATG CAA AAT TTA GAA Tyr Leu Gln Ser Ala Pro Leu Gln Leu Met Ser Met Gln Asn Leu Glu 25 30 35	449

					_							-			ATT Ile 55	497
														TGT Cys 70		545
														AGC Ser		593
														GCT Ala		641
														GAA Glu		689
														CTG Leu		737
														GAA Glu 150		785
														GAA Glu		833
														TTG Leu		881
														GAA Glu		929
														GCT Ala		977
														GGC Gly 230		1025
														GAG Glu		1073
														GGA Gly		1121
TCA Ser	CCC Pro 265	AAA Lys	AGG Arg	CGT Arg	CGT Arg	CTC Leu 270	AGC Ser	TCG Ser	TCT Ser	CTA Leu	AAC Asn 275	CCT Pro	TCT Ser	AAA Lys	AGA Arg	1169
GCA Ala 280	CCA Pro	AAA Lys	CAG Gln	ACT Thr	GAG Glu 285	GAA Glu	ATT Ile	AAA Lys	CAT His	GTG Val 290	GAC Asp	ATG Met	AAC Asn	CAA Gln	AAG Lys 295	1217
														CAG Gln 310		1265

TCC Ser	CTT Leu	GAA Glu	TAC Tyr 315	AGT Ser	GGC Gly	CTA Leu	AAG Lys	AAT Asn 320	CCT Pro	GTT Val	ATT Ile	GAG Glu	ATG Met 325	TTA Leu	GAA Glu	1313
GGA Gly	ATT Ile	GCT Ala 330	GTT Val	GTC Val	TTA Leu	CAA Gln	CTG Leu 335	ACT Thr	GCT Ala	CTG Leu	TGT Cys	ACT Thr 340	GTT Val	CAT His	TGT Cys	1361
TCT Ser	CAT His 345	CAA Gln	AAC Asn	ATG Met	AAC Asn	TGC Cys 350	CGT Arg	ACT Thr	TTC Phe	AAG Lys	GAC Asp 355	TGT Cys	CAA Gln	CAT His	AAA Lys	1409
					TCT Ser 365											1457
TAC Tyr	ACA Thr	AAA Lys	GTG Val	CTT Leu 380	AAG Lys	AGC Ser	TGT Cys	AGA Arg	AGT Ser 385	TTG Leu	TTA Leu	GAA Glu	TCT Ser	GTT Val 390	CAG Gln	1505
					GCA Ala											1553
GCT Ala	TTG Leu	ATT Ile 410	TAT Tyr	ATG Met	CAA Gln	GTA Val	AAC Asn 415	AGT Ser	TCA Ser	TTT Phe	GAA Glu	GAT Asp 420	CAT His	ATC Ile	CTG Leu	1601
					ATG Met											1649
					AAG Lys 445	Leu										1697
					TCA Ser											1745
					ACT Thr											1793
					AAC Asn											1841
					AGT Ser											1889
					CCC Pro 525											1937
TCT Ser	GAC Asp	ATT	GTC Val	AAG Lys 540	AAA Lys	GAA Glu	TTT Phe	GCT Ala	TCT Ser 545	ATA Ile	CTT Leu	GGT Gly	CAA Gln	CTT Leu 550	GTC Val	1985
TGT Cys	ACT Thr	CTT Leu	CAC His 555	GGC Gly	ATG Met	TTT Phe	TAT Tyr	CTG Leu 560	ACA Thr	AGT Ser	TCT Ser	TTA Leu	ACA Thr 565	GAA Glu	CCT Pro	2033
TTC Phe	TCT	GAA Glu 570	CAC His	GGA Gly	CAT His	GTG Val	GAC Asp 575	CTC Leu	TTC Phe	TGT Cys	AGG Arg	AAC Asn 580	TTG Leu	AAA Lys	GCC Ala	2081

															TGC Cys	2129
AAG Lys 600	CCA Pro	TTC Phe	CTT Leu	TTC Phe	CTA Leu 605	CTG Leu	AAA Lys	AAA Lys	AAA Lys	ATA Ile 610	CCT Pro	AGT Ser	CCA	GTA Val	AAA Lys 615	2177
						CTA Leu										2225
						GTA Val										2273
TTA Leu	ATG Met	GAA Glu 650	GAT Asp	CCA Pro	GAC Asp	AAA Lys	GAT Asp 655	GTT Val	AGA Arg	GTG Val	GCT Ala	TTT Phe 660	AGT Ser	GGA Gly	TAA naA	2321
						TCC Ser 670										2369
GAG Glu 680	CTT	TTT Phe	GTC Val	TTA Leu	AGA Arg 685	ATG Met	AAG Lys	GAA Glu	GCA Ala	TAT Tyr 690	ACA Thr	CAT His	GCC Ala	CAA Gln	ATA Ile 695	2417
TCA Ser	AGA Arg	AAT Asn	AAT Asn	GAG Glu 700	CTG Leu	AAG Lys	GAT Asp	ACC Thr	TTG Leu 705	ATT Ile	CTT Leu	ACA Thr	ACA Thr	GGG Gly 710	GAT Asp	2465
ATT Ile	GGA Gly	AGG Arg	GCC Ala 715	GCA Ala	AAA Lys	GGA Gly	GAT Asp	TTG Leu 720	GTA Val	CCA Pro	TTT Phe	GCA Ala	CTC Leu 725	TTA Leu	CAC His	2513
						TCC Ser										2561
TAC Tyr	ACA Thr 745	GAA Glu	ATT Ile	AGA Arg	GCT Ala	CTG Leu 750	GTT Val	GCA Ala	GCT Ala	AAA Lys	AGT Ser 755	GTT Val	AAA Lys	CTG Leu	CAA Gln	2609
AGT Ser 760	TTT Phe	TTC Phe	AGC Ser	CAG Gln	TAT Tyr 765	AAG Lys	AAA Lys	CCC Pro	ATC Ile	TGT Cys 770	CAG Gln	TTT Phe	TTG Leu	GTA Val	GAA Glu 775	2657
TCC Ser	CTT Leu	CAC His	TCT Ser	AGT Ser 780	CAG Gln	ATG Met	ACA Thr	GCA Ala	CTT Leu 785	CCG Pro	AAT Asn	ACT Thr	CCA Pro	TGC Cys 790	CAG Gln	2705
AAT Asn	GCT Ala	GAC Asp	GTG Val 795	CGA Arg	AAA Lys	CAA Gln	GAT Asp	GTG Val 800	GCT Ala	CAC His	CAG Gln	AGA Arg	GAA Glu 805	ATG Met	GCT Ala	2753
TTA Leu	AAT Asn	ACG Thr 810	TTG Leu	TCT Ser	GAA Glu	ATT Ile	GCC Ala 815	AAC Asn	GTT Val	TTC Phe	GAC Asp	TTT Phe 820	CCT Pro	GAT Asp	CTT Leu	2801
AAT Asn	CGT Arg 825	TTT Phe	CTT Leu	ACT Thr	AGG Arg	ACA Thr 830	TTA Leu	CAA Gln	GTT Val	CTA Leu	CTA Leu 835	CCT Pro	GAT Asp	CTT Leu	GCT Ala	2849
GCC Ala 840	AAA Lys	GCA Ala	AGC Ser	CCT Pro	GCA Ala 845	GCT Ala	TCT Ser	GCT Ala	CTC Leu	ATT Ile 850	CGA Arg	ACT Thr	TTA Leu	GGA Gly	AAA Lys 855	2897

				A AAC AAC TTO e Asn Asn Pho		2945
		. Cys Ser C		A GAT GAA TTA s Asp Glu Leu 885	ı Glu Arg	2993
	Tyr Leu Lys			T GAA CTG GGO e Glu Leu Gly 900		3041
				A TTA TTG CTC u Leu Leu Leu 915		3089
		Val Phe A		G TCA ATA CTT u Ser Ile Leu 0		3137
TTT GCA TCC Phe Ala Ser	AGT GAT GAT Ser Asp Asp 940	CCA TAT C. Pro Tyr G	CAG GGC CCG Sln Gly Pro 945	G AGA GAT ATO O Arg Asp Ile	ATA TCA Ile Ser 950	3185
		Tyr Leu G		A TTG TTG GGC s Leu Leu Gly 965	· Ile Leu	3233
				T GTT GGC ATT r Val Gly Ile 980		3281
				G ATG AAG TTA u Met Lys Leu 995		3329
		Val Arg V		G ATG ACC ACA t Met Thr Thr 10		3377
				A TTG TGT TGC u Leu Cys Cys		3425
		Cys Leu A		r TGT CTG GGC a Cys Leu Gly 104	Ser Leu	3473
	Val Ile Val			T ATA CAC ATO u Ile His Ile 1060		3521
				A ATT GAA AAC e Ile Glu Asn 1075		3569
		His Glu I		T TTA CCT GAT e Leu Pro Asp 90		3617
				A TAC AGA AAG u Tyr Arg Lys		3665
		Gln Thr T		G CTC TCT ATG Leu Ser Met 112	Lys Ala	3713

ATT Ile	CAA Gln	CAT His	GAA Glu O	AAT Asn	GTC Val	GAT Asp	GTT Val 113	Arg	ATT Ile	CAT His	GCT Ala	CTT Leu 114	Thr	AGC Ser	TTG Leu	3	761
		Thr	TTG Leu				Gln					Lys				3	809
	Ser		ACA Thr			Pro					Leu					3	857 [°]
			TGC Cys		Asp					Ala					Gly	3:	905
			GGG Gly 1199	Glu					Asp					Asp		3	953
			GAA Glu O					Asp					Thr			4	001
		Ser	AGC Ser				Gly					Leu				4	049
	Leu		TAT Tyr			Asn					Asp					47	097
			GAG Glu		Leu					Cys					Thr	4	145
			GGT Gly 1275	His					Arg					Val		4:	193
			GAA Glu					Thr					Ser			4:	241
		Asp	TGG Trp				Lys					Leu				4:	289
	Ser		TTT Phe			${\tt Trp}$					Ala					4:	337
			CGA Arg		Asp					Ile					Ser	4:	385
ATT Ile	ATG Met	ATG Met	AAG Lys 1355	His	GAT Asp	TTC Phe	AAA Lys	GTG Val 1360	Thr	ATC Ile	TAT Tyr	CTT Leu	CTT Leu 1365	Pro	CAT His	4.4	433
ATT Ile	CTG Leu	GTG Val 1370	TAT Tyr)	GTC Val	TTA Leu	CTG Leu	GGT Gly 1375	Cys	AAT Asn	CAA Gln	GAA Glu	GAT Asp 1380	Gln	CAG Gln	GAG Glu	4.4	481
GTT Val	TAT Tyr 1385	Ala	GAA Glu	ATT Ile	ATG Met	GCA Ala 1390	Val	CTA Leu	AAG Lys	CAT His	GAC Asp 1395	Asp	CAG Gln	CAT His	ACC Thr	45	529

-154-

140	Asr.	ı Tnr	Glr	Asp	140	Ala 5	Ser	as T) Leu	Cys 141	Gln 0	Leu	Sei	Thr	CAG Gln 1415	4	4577
1111	Val	Pne	e ser	142	0 Leu	Asp	His	Lei	142	Gln 5	Trp	Ala	Arç	His 143	•	. 4	1625
Pile	GIN	Ala	143	Lys 5	Ala	Glu	Lys	Cys 144	Pro	His	Ser	Lys	Ser 144	Asn 5	AGA Arg	4	1673
Asn	Lys	Val 145	Asp 0	Ser	Met	Val	Ser 145	Thr 5	· Val	Asp	Tyr	Glu 146	Asp 0	Tyr		4	721
ser	146	onr 5	Arg	Pne	Leu	147	Leu 0	Ile	CCC Pro	Gln	Asp 147	Thr 5	Leu	Ala	Val	4	769
148	Ser O	Pne	Arg	Ser	Lys 148	Ala 5	Tyr	Thr	CGA Arg	Ala 1490	Val	Met	His	Phe	Glu 1495	4	817
TCA Ser	TTT	ATT	ACA Thr	GAA Glu 150	Lys	AAG Lys	CAA Gln	AAT Asn	ATT Ile 1509	Gln	GAA Glu	CAT His	CTT Leu	GGA Gly 151	Phe	4	865
TTA Leu	CAG Gln	AAA Lys	TTG Leu 151	Tyr	GCT Ala	GCT Ala	ATG Met	CAT His 152	GAA Glu O	CCT Pro	GAT Asp	GGA Gly	GTG Val 152	Ala	GGA Gly	4	913
GTC Val	AGT Ser	GCA Ala 1530	Ile	AGA Arg	AAG Lys	GCA Ala	GAA Glu 1535	Pro	TCT Ser	CTA Leu	AAA Lys	GAA Glu 1540	Gln	ATC Ile	CTT Leu	4	961
GAA Glu	CAT His 1545	Glu	AGC Ser	CTT Leu	GGC Gly	TTG Leu 1550	Leu	AGG Arg	GAT Asp	GCC Ala	ACT Thr 1555	Ala	TGT Cys	TAT Tyr	GAC Asp	5	009
AGG Arg 1560	Ala	ATT Ile	CAG Gln	CTA Leu	GAA Glu 1565	Pro	GAC Asp	CAG Gln	ATC Ile	ATT Ile 1570	His	TAC Tyr	CAT His	GGT Gly	GTA Val 1575	5(057
GTA Val	AAG Lys	TCC Ser	ATG Met	TTA Leu 1580	Gly	CTT Leu	GGT Gly	CAG Gln	CTG Leu 1585	Ser	ACT Thr	GTT Val	ATC Ile	ACT Thr 1590	Gln	5	105
GTG Val	AAT Asn	GGA Gly	GTG Val 1595	His	GCT Ala	AAC Asn	AGG Arg	TCC Ser 1600	GAG Glu)	TGG Trp	ACA Thr	GAT Asp	GAA Glu 1605	Leu	AAC Asn	52	153
ACG Thr	TAC Tyr	AGA Arg 1610	vaı	GAA Glu	GCA Ala	GCT Ala	TGG Trp 1615	Lys	TTG Leu	TCA Ser	CAG Gln	TGG Trp 1620	Asp	TTG Leu	GTG Val	52	201
GAA Glu	AAC Asn 1625	lyr	TTG Leu	GCA Ala	GCA Ala	GAT Asp 1630	Gly	AAA Lys	TCT Ser	Thr	ACA Thr 1635	\mathtt{Trp}	AGT Ser	GTC Val	AGA Arg	52	249
CTG Leu 1640	GIA	CAG Gln	CTA Leu	Leu	TTA Leu 1645	Ser	GCC Ala	AAA Lys	AAA Lys	AGA Arg 1650	Asp	ATC Ile	ACA Thr	GCT Ala	TTT Phe 1655	52	297
TAT Tyr	GAC Asp	TCA Ser	Leu	AAA Lys 1660	CTA Leu	GTG Val	AGA Arg	GCA Ala	GAA Glu 1665	CAA . Gln	ATT Ile	GTA Val	CCT Pro	CTT Leu 1670	Ser	53	345

GCT GCA AGC Ala Ala Ser	TTT GAA AGA Phe Glu Arg 1675	Gly Ser T	TAC CAA CGA Tyr Gln Arg 1680	GGA TAT GAA Gly Tyr Glu 1685	Tyr Ile
	His Met Leu			AGC ATC AAA Ser Ile Lys 1700	
				GAT TCT CTA Asp Ser Leu 1715	
		Thr Gln A		AGA GCC AAG Arg Ala Lys	
ATC CTG GCT Ile Leu Ala	CTC CGG AGG Leu Arg Arg 1740	GCT TTA C Ala Leu L	CTA AGC CTC Leu Ser Leu 1745	AAC AAA AGA Asn Lys Arg	CCA GAT 5589 Pro Asp 1750
		Glu Cys T		AGT GCC AGG Ser Ala Arg 1765	Val Ala
	Gly His His			GCT CTC CTT Ala Leu Leu 1780	
				AGG GCA AAG Arg Ala Lys 1795	
		His Gln A		GTT CTT CAA Val Leu Gln	
				CCT GAG GGT Pro Glu Gly	
		Ala Met L		GGC CGA TTT Gly Arg Phe 1845	Met Glu
	Asn Phe Glu			AAA AAA TAT Lys Lys Tyr 1860	
				CAT TTT TAC His Phe Tyr 1875	
		Met Pro M		GAC AAC AAA Asp Asn Lys	
				CAT TTT GGC . His Phe Gly .	
		Phe Ile T		ATG CCA CGA Met Pro Arg 1925	
	Leu Asp Tyr			GAA TGG GAA . Glu Trp Glu 1940	

GGC CGC TCC Gly Arg Ser 1945	GAT CGT GTA Asp Arg Val	CAA ATG AG Gln Met Ar 1950	G AAT GAT g Asn Asp	TTG GGT AAA Leu Gly Lys 1955	ATA AAC 6209 Ile Asn
AAG GTT ATC Lys Val Ile 1960	ACA GAG CAT Thr Glu His 196	Thr Asn Ty	T TTA GCT r Leu Ala 1970	Pro Tyr Gln	TTT TTG 625° Phe Leu 1975
ACT GCT TTT Thr Ala Phe	TCA CAA TTG Ser Gln Leu 1980	ATC TCT CG. Ile Ser Ar	A ATT TGT g Ile Cys 1985	CAT TCT CAC His Ser His	GAT GAA 6305 Asp Glu 1990
GTT TTT GTT Val Phe Val	GTC TTG ATG Val Leu Met 1995	GAA ATA AT. Glu Ile Ile 20	e Ala Lys	GTA TTT CTA Val Phe Leu 2005	Ala Tyr
CCT CAA CAA Pro Gln Gln 201	GCA ATG TGG Ala Met Trp 0	ATG ATG AC Met Met Th 2015	A GCT GTG '	TCA AAG TCA Ser Lys Ser 2020	TCT TAT 6401 Ser Tyr
CCC ATG CGT Pro Met Arg 2025	GTG AAC AGA Val Asn Arg	TGC AAG GA Cys Lys Gl 2030	u Ile Leu i	AAT AAA GCT Asn Lys Ala 2035	ATT CAT 6449 Ile His
ATG AAA AAA Met Lys Lys 2040	TCC TTA GAG Ser Leu Glu 204	Lys Phe Val	T GGA GAT (1 Gly Asp 2 2050	GCA ACT CGC Ala Thr Arg	CTA ACA 6497 Leu Thr 2055
GAT AAG CTT Asp Lys Leu	CTA GAA TTG Leu Glu Leu 2060	TGC AAT AAI Cys Asn Lys	A CCG GTG (s Pro Val (2065	Glu Ile Leu .	GCT TCT 6545 Ala Ser 2070
CTT CAG AAA Leu Gln Lys	CCA AAG AAG Pro Lys Lys 2075	ATT TCT TT	u Lys Gly :	TCA GAT GGA . Ser Asp Gly 2085	AAG TTC 6593 Lys Phe
TAC ATC ATG Tyr Ile Met 209	ATG TGT AAG Met Cys Lys 0	CCA AAA GA Pro Lys Asi 2095	T GAC CTG I	AGA AAG GAT Arg Lys Asp 2100	TGT AGA 6641 Cys Arg
CTA ATG GAA Leu Met Glu 2105	TTC AAT TCC Phe Asn Ser	TTG ATT AAT Leu Ile Ass 2110	n Lys Cys 1	TTA AGA AAA (Leu Arg Lys : 2115	GAT GCA 6689 Asp Ala
GAG TCT CGT Glu Ser Arg 2120	AGA AGA GAA Arg Arg Glu 212	Leu His Ile	T CGA ACA T e Arg Thr 7 2130	TAT GCA GTT . Tyr Ala Val	ATT CCA 6737 Ile Pro 2135
CTA AAT GAT Leu Asn Asp	GAA TGT GGG Glu Cys Gly 2140	ATT ATT GAZ	A TGG GTG I u Trp Val I 2145	Asn Asn Thr .	GCT GGT 6785 Ala Gly 2150
TTG AGA CCT Leu Arg Pro	ATT CTG ACC Ile Leu Thr 2155	AAA CTA TAT Lys Leu Tyr 216	r Lys Glu 1	AAG GGA GTG ' Lys Gly Val ' 2165	TAT ATG 6833 Tyr Met
ACA GGA AAA Thr Gly Lys 2170	GAA CTT CGC Glu Leu Arg 0	CAG TGT ATO Gln Cys Met 2175	G CTA CCA A t Leu Pro I	AAG TCA GCA (Lys Ser Ala . 2180	GCT TTA 6881 Ala Leu
TCT GAA AAA Ser Glu Lys 2185	CTC AAA GTA Leu Lys Val	TTC CGA GAA Phe Arg Glu 2190	u Phe Leu I	CTG CCC AGG (Leu Pro Arg 1 2195	CAT CCT 6929 His Pro
CCT ATT TTT Pro Ile Phe 2200	CAT GAG TGG His Glu Trp 2209	Phe Leu Arg	A ACA TTC (g Thr Phe I 2210	CCT GAT CCT A	ACA TCA 6977 Thr Ser 2215

-157-

			AGT Ser		Ser					Ser					Ser		7025
			TAT Tyr 2235	Ile					Asp					Asn			7073
			TCT Ser					Cys					Phe				7121
		Asn	AAG Lys				Phe					Ile					7169
	Leu		CAT His			Val					Pro					•	721 7
			CGA Arg		Ala					Met					Asp	•	7265
			CCT Pro 2315	Leu					Lys					Asp			7313
			TGG Trp					Lys					Ala			-	7361
		Thr	GGA Gly				Asn					Thr				•	7409
	Ile		CAG Gln			Gln					Thr					-	7457
ACA Thr	GGA Gly	CTG Leu	CCG Pro	TTA Leu 2380	Ser	ATT Ile	GAA Glu	GGA Gly	CAT His 2385	Val	CAT His	TAC Tyr	CTT Leu	ATA Ile 2390	Gln	7	7505
GAA Glu	GCT Ala	ACT Thr	GAT Asp 2395	Glu	AAC Asn	TTA Leu	CTA Leu	TGC Cys 2400	Gln	ATG Met	TAT Tyr	CTT Leu	GGT Gly 2405	Trp	ACT Thr	7	7553
	TAT Tyr		TGAA	ATGA	T AA.	TATG	AAAT	A GA	rata.	GTTA	ATA	ATCT	AAA			7	7602
AGTA	AAAA	AA A	AAAA	AAAA	AA AA											7	7624

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2410 amino acids
 (B) TYPE: amino acid

 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

WO 97/18323

1	GIY	nis	Ala	5	GIU	пр	PIO	Vall	10	Met	Sei	Arg	Pne	15	Sei
Gln	Leu	Asp	Glu 20	His	Met	Gly	Tyr	Leu 25	Gln	Ser	Ala	Pro	Leu 30	Gln	Let
Met	Ser	Met 35	Gln	Asn	Leu	Glu	Phe 40	Ile	Glu	Val	Thr	Leu 45	Leu	Met	Val
Leu	Thr 50	Arg	Ile	Ile	Ala	Ile 55	Val	Phe	Phe	Arg	Arg 60	Gln	Glu	Leu	Leu
Leu 65	Trp	Gln	Ile	Gly	Cys 70	Val	Leu	Leu	Glu	Tyr 75	Gly	Ser	Pro	Lys	11e 80
Lys	Ser	Leu	Ala	Ile 85	Ser	Phe	Leu	Thr	Glu 90	Leu	Phe	Gln	Leu	Gly 95	Gly
Leu	Pro	Ala	Gln 100	Pro	Ala	Ser	Thr	Phe 105	Phe	Ser	Ser	Phe	Leu 110	Glu	Leu
Leu	Lys	His 115	Leu	Val	Glu	Met	Asp 120	Thr	Asp	Gln	Leu	Lys 125	Leu	Tyr	Glu
Glu	Pro 130	Leu	Ser	Lys	Leu	Ile 135	Lys	Thr	Leu	Phe	Pro 140	Phe	Glu	Ala	Glu
Ala 145	Tyr	Arg	Asn	Ile	Glu 150	Pro	Val	Tyr	Leu	Asn 155	Met	Leu	Leu	Glu	Lys 160
Leu	Cys	Val	Met	Phe 165	Glu	Asp	Gly	Val	Leu 170	Met	Arg	Leu	Lys	Ser 175	Asp
Leu	Leu	Lys	Ala 180	Ala	Leu	Cys	His	Leu 185	Leu	Gln	Tyr	Phe	Leu 190	Lys	Phe
Val	Pro	Ala 195	Gly	Tyr	Glu	Ser	Ala 200	Leu	Gln	Val	Arg	Lys 205	Val	Tyr	Val
Arg	Asn 210	Ile	Cys	Lys	Ala	Leu 215	Leu	Asp	Val	Leu	Gly 220	Ile	Glu	Val	Asp
Ala 225	Glu	Tyr	Leu	Leu	Gly 230	Pro	Leu	Tyr	Ala	Ala 235	Leu	Lys	Met	Glu	Ser 240
Met	Glu	Ile	Ile	Glu 245	Glu	Ile	Gln	Cys	Gln 250	Thr	Gln	Gln	Glu	Asn 255	Leu
Ser	Ser	Asn	Ser 260	Asp	Gly	Ile	Ser	Pro 265	Lys	Arg	Arg	Arg	Leu 270	Ser	Ser
Ser	Leu	Asn 275	Pro	Ser	Lys	Arg	Ala 280	Pro	Lys	Gln	Thr	Glu 285	Glu	Ile	Lys
His	Val 290	Asp	Met	Asn	Gln	Lys 295	Ser	Ile	Leu	Trp	Ser 300	Ala	Leu	Lys	Glr
Lys 305	Ala	Glu	Ser	Leu	Gln 310	Ile	Ser	Leu	Glu	Tyr 315	Ser	Gly	Leu	Lys	Asr. 320
Pro	Val	Ile	Glu	Met 325	Leu	Glu	Gly	Ile	Ala 330	Val	Val	Leu	Gln	Leu 335	Thr
Ala	Leu	Cys	Thr 340	Val	His	Cys	Ser	His 345	Gln	Asn	Met	Asn	Cys 350	Arg	Thr

Phe	Lys	Asp 355	Cys	Gln	His	Lys	Ser 360	Lys	Lys	Lys	Pro	Ser 365	Val	Val	Ile	
Thr	Trp 370	Met	Ser	Leu	Asp	Phe 375	Tyr	Thr	Lys	Val	Leu 380	Lys	Ser	Cys	Arg	
Ser 385	Leu	Leu	Glu	Ser	Val 390	Gln	Lys	Leu	Asp	Leu 395	Glu	Ala	Thr	Ile	Asp 400	
Lys	Val	Val	Lys	Ile 405	Tyr	Asp	Ala	Leu	Ile 410	Tyr	Met	Gln	Val	Asn 415	Ser	
Ser	Phe	Glu	Asp 420	His	Ile	Leu	Glu	Asp 425	Leu	Cys	Gly	Met	Leu 430	Ser	Leu	
Pro	Trp	Ile 435	Tyr	Ser	His	Ser	Asp 440	Asp	Gly	Cys	Leu	Lys 445	Leu	Thr	Thr	
Phe	Ala 450	Ala	Asn	Leu	Leu	Thr 455	Leu	Ser	Cys	Arg	Ile 460	Ser	Asp	Ser	Tyr	
Ser 465	Pro	Gln	Ala	Gln	Ser 470	Arg	Cys	Val	Phe	Leu 475	Leu	Thr	Leu	Phe	Pro 480	
Arg	Arg	Ile	Phe	Leu 485	Glu	Trp	Arg	Thr	Ala 490	Val	Tyr	Asn	Trp	Ala 495	Leu	
Gln	Ser	Ser	His 500	Glu	Val	Ile	Arg	Ala 505	Ser	Cys	Val	Ser	Gly 510	Phe	Phe	
Ile	Leu	Leu 515	Gln	Gln	Gln	Asn	Ser 520	Cys	Asn	Arg	Val	Pro 525	Lys	Ile	Leu	
Ile	Asp 530	Lys	Val	Lys	Asp	Asp 535	Ser	Asp	Ile	Val	Lys 540	Lys	Glu	Phe	Ala	
Ser 545	Ile	Leu	Gly	Gln	Leu 550	Val	Cys	Thr	Leu	His 555	Gly	Met	Phe	Tyr	Leu 560	
Thr	Ser	Ser	Leu	Thr 565	Glu	Pro	Phe	Ser	Glu 570	His	Gly	His	Val	Asp 575	Leu	
Phe	Cys	Arg	Asn 580	Leu	Lys	Ala	Thr	Ser 585	Gln	His	Glu	Cys	Ser 590	Ser	Ser '	
Gln	Leu	Lys 595	Ala	Ser	Val	Cys	Lys 600	Pro	Phe	Leu	Phe	Leu 605	Leu	Lys	Lys	
Lys	Ile 610	Pro	Ser	Pro	Val	Lys 615	Leu	Ala	Phe	Ile	Asp 620	Asn	Leu	His	His	
Leu 625	Cys	Lys	His	Leu	Asp 630	Phe	Arg	Glu	Asp	Glu 635	Thr	Asp	Val	Lys	Ala 640	
Val	Leu	Gly	Thr	Leu 645	Leu	Asn	Leu	Met	Glu 650	Asp	Pro	Asp	Lys	Asp 655	Val	
Arg	Val	Ala	Phe 660	Ser	Gly	Asn	Ile	Lys 665	His	Ile	Leu	Glu	Ser 670	Leu	Asp	
Ser	Glu	Asp 675	Gly	Phe	Ile	Lys	Glu 680	Leu	Phe	Val	Leu	Arg 685	Met	Lys	Glu	
Ala	Tyr 690	Thr	His	Ala	Gln	Ile 695	Ser	Arg	Asn	Asn	Glu 700	Leu	Lys	Asp	Thr	

Leu 705	Ile	Leu	Thr	Thr	Gly 710	Asp	Ile	Gly	Arg	Ala 715	Ala	Lys	Gly	Asp	Leu 720
Val	Pro	Phe	Ala	Leu 725	Leu	His	Leu	Leu	His 730	Cys	Leu	Leu	Ser	Lys 735	Ser
Ala	Ser	Val	Ser 740	Gly	Ala	Ala	Tyr	Thr 745	Glu	Ile	Arg	Ala	Leu 750	Val	Ala
Ala	Lys	Ser 755	Val	Lys	Leu	Gln	Ser 760	Phe	Phe	Ser	Gln	Tyr 765	Lys	Lys	Pro
Ile	Cys 770	Gln	Phe	Leu	Val	Glu 775	Ser	Leu	His	Ser	Ser 780	Gln	Met	Thr	Ala
Leu 785	Pro	Asn	Thr	Pro	Cys 790	Gln	Asn	Ala	Asp	Val 795	Arg	Lys	Gln	Asp	Val 800
Ala	His	Gln	Arg	Glu 805	Met	Ala	Leu	Asn	Thr 810	Leu	Ser	Glu	Ile	Ala 815	Asn
Val	Phe	Asp	Phe 820	Pro	Asp	Leu	Asn	Arg 825	Phe	Leu	Thr	Arg	Thr 830	Leu	Gln
Val	Leu	Leu 835	Pro	Asp	Leu	Ala	Ala 840	Lys	Ala	Ser	Pro	Ala 845	Ala	Ser	Ala
Leu	Ile 850	Arg	Thr	Leu	Gly	Lys 855	Gln	Leu	Asn	Val	Asn 860	Arg	Arg	Glu	Ile
Leu 865	Ile	Asn	Asn	Phe	Lys 870	Tyr	Ile	Phe	Ser	His 875	Leu	Val	Cys	Ser	ay5
Ser	Lys	Asp	Glu	Leu 885	Glu	Arg	Ala	Leu	His 890	Tyr	Leu	Lys	Asn	Glu 895	Thr
Glu	Ile	Glu	Leu 900	Gly	Ser	Leu	Leu	Arg 905	Gln	Asp	Phe	Gln	Gly 910	Leu	His
Asn	Glu	Leu 915	Leu	Leu	Arg	Ile	Gly 920	Glu	His	Tyr	Gln	Gln 925	Val	Phe	Asn
Gly	Leu 930	Ser	Ile	Leu	Ala	Ser 935	Phe	Ala	Ser	Ser	Asp 940	Asp	Pro	Tyr	Gln
Gly 945	Pro	Arg	Asp	Ile	Ile 950	Ser	Pro	Glu	Leu	Met 955	Ala	qaA	Tyr	Leu	Gln 960
Pro	Lys	Leu	Leu	Gly 965	Ile	Leu	Ala	Phe	Phe 970	Asn	Met	Gln	Leu	Leu 975	Ser
Ser	Ser	Val	Gly 980	Ile	Glu	qäA	Lys	Lys 985	Met	Ala	Leu	Asn	Ser 990	Leu	Met
Ser	Leu	Met 995	Lys	Leu	Met	Gly	Pro 1000		His	Val	Ser	Ser 1005		Arg	Val
Lys	Met 1010	Met	Thr	Thr	Leu	Arg 1015		Gly	Leu	Arg	Phe 1020		Asp	qaA	Phe
Pro 1025	Glu	Leu	Cys	Cys	Arg 1030		Trp	Asp	Cys	Phe 1035		Arg	Cys	Leu	Asp 1040
His	Ala	Cys	Leu	Gly 1045	Ser	Leu	Leu	Ser	His 1050		Ile	Val	Ala	Leu 1055	

-161-

- Pro Leu Ile His Ile Gln Pro Lys Glu Thr Ala Ala Ile Phe His Tyr 1060 1065 1070
- Leu Ile Ile Glu Asn Arg Asp Ala Val Gln Asp Phe Leu His Glu Ile 1075 1080 1085
- Tyr Phe Leu Pro Asp His Pro Glu Leu Lys Lys Ile Lys Ala Val Leu 1090 1095 1100
- Gln Glu Tyr Arg Lys Glu Thr Ser Glu Ser Thr Asp Leu Gln Thr Thr 1105 1110 1115
- Leu Gln Leu Ser Met Lys Ala Ile Gln His Glu Asn Val Asp Val Arg 1125 1130 1135
- Ile His Ala Leu Thr Ser Leu Lys Glu Thr Leu Tyr Lys Asn Glu Glu 1140 1145 1150
- Lys Leu Ile Lys Tyr Ala Thr Asp Ser Glu Thr Val Glu Pro Ile Ile 1155 1160 1165
- Ser Gln Leu Val Thr Val Leu Leu Lys Gly Cys Gln Asp Ala Asn Ser 1170 1180
- Gln Ala Arg Leu Leu Cys Gly Glu Cys Leu Gly Glu Leu Gly Ala Ile 1185 1190 1195 1200
- Asp Pro Gly Arg Leu Asp Phe Ser Thr Thr Glu Thr Gln Gly Lys Asp 1205 1210 1215
- Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu 1220 1225 1230
- Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg 1235 1240 1245
- Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr 1250 1260
- Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg 1265 1270 1275 1280
- Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr 1285 1290 1295
- Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys 1300 1310
- Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala 1315 1320 1325
- Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1330 1340
- Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val 1345 1350 1355 1360
- Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1365 1370 1375
- Asn Gln Glu Asp Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 1385 1390
- Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395

- Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 1415 1420
- Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 1430 1435 1440
- Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 1450 1455
- Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 1465 1470
- Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 1480 1485
- Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 1495 1500
- Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 1510 1515 1520
- Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 1530 1535
- Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 1545 1550
- Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 1560 1565
- Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1570 1575 1580
- Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser 1585 1590 1595 1600
- Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys 1605 1610 1615
- Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys 1620 1625 1630
- Ser Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Ser Ala Lys 1635 1640 1645
- Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala 1650 1655 1660
- Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr 1665 1670 1680
- Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu 1685 1690 1695
- Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser 1700 1705 1710
- Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn 1715 1720 1725
- Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu 1730 1740
- Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp 1745 1750 1755 1760

WO 97/18323 PCT/US96/19337

-163-

Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala 1765 1770 1775

Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr 1780 1785 1790

Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala 1795 1800 1805

Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu 1810 1815 1820

Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu 1825 1830 1835 1840

Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala 1845 1850 1855

Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu 1860 1865 1870

Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met 1875 1880 1885

Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile 1890 1895 1900

Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr 1905 1910 1915 1920

Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys 1925 1930 1935

Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg 1940 1945 1950

Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr 1955 1960 1965

Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg 1970 1975 1980

Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Met Glu Ile Ile 1985 1990 1995 2000

Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr 2005 2010 2015

Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu 2020 2025 2030

Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val 2035 2040 2045

Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys 2050 2060

Pro Val Glu Ile Leu Ala Ser Leu Gln Lys Pro Lys Lys Ile Ser Leu 2065 2070 2075

Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met Met Cys Lys Pro Lys Asp 2085 2090 2095

Asp Leu Arg Lys Asp Cys Arg Leu Met Glu Phe Asn Ser Leu Ile Asn 2100 2105 2110

- Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg Arg Arg Glu Leu His Ile
- Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp Glu Cys Gly Ile Ile Glu 2135
- Trp Val Asn Asn Thr Ala Gly Leu Arg Pro Ile Leu Thr Lys Leu Tyr 2150 2155
- Lys Glu Lys Gly Val Tyr Met Thr Gly Lys Glu Leu Arg Gln Cys Met 2165
- Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys Leu Lys Val Phe Arg Glu 2185
- Phe Leu Leu Pro Arg His Pro Pro Ile Phe His Glu Trp Phe Leu Arg 2200 2205
- Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser Ser Arg Ser Ala Tyr Cys 2215
- Arg Ser Thr Ala Val Met Ser Met Val Gly Tyr Ile Leu Gly Leu Gly 2230
- Asp Arg His Gly Glu Asn Ile Leu Phe Asp Ser Leu Thr Gly Glu Cys 2245 2250
- Val His Val Asp Phe Asn Cys Leu Phe Asn Lys Gly Glu Thr Phe Glu 2265
- Val Pro Glu Ile Val Pro Phe Arg Leu Thr His Asn Met Val Asn Gly
- Met Gly Pro Met Gly Thr Glu Gly Leu Phe Arg Arg Ala Cys Glu Val 2295
- Thr Met Arg Leu Met Arg Asp Gln Arg Glu Pro Leu Met Ser Val Leu 2310 2315
- Lys Thr Phe Leu His Asp Pro Leu Val Glu Trp Ser Lys Pro Val Lys
- Gly His Ser Lys Ala Pro Leu Asn Glu Thr Gly Glu Val Val Asn Glu 2340 2345
- Lys Ala Lys Thr His Val Leu Asp Ile Glu Gln Arg Leu Gln Gly Val 2360
- Ile Lys Thr Arg Asn Arg Val Thr Gly Leu Pro Leu Ser Ile Glu Gly 2375
- His Val His Tyr Leu Ile Gln Glu Ala Thr Asp Glu Asn Leu Leu Cys 2395
- Gln Met Tyr Leu Gly Trp Thr Pro Tyr Met 2405

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7502 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

 - (D) TOPOLOGY: linear

WO 97/18323 PCT/US96/19337 -

-165-

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS (B) LOCATION: 1..7440

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATG Met 1	GGT Gly	CAT His	GCT Ala	GTG Val 5	GAA Glu	TGG Trp	CCA Pro	GTG Val	GTC Val 10	ATG Met	AGC Ser	CGA Arg	TTT Phe	TTA Leu 15	AGT Ser	48
CAA Gln	TTA Leu	GAT Asp	GAA Glu 20	CAC His	ATG Met	GGA Gly	TAT Tyr	TTA Leu 25	CAA Gln	TCA Ser	GCT Ala	CCT Pro	TTG Leu 30	CAG Gln	TTG Leu	96
ATG Met	AGT Ser	ATG Met 35	CAA Gln	AAT Asn	TTA Leu	GAA Glu	TTT Phe 40	ATT Ile	GAA Glu	GTC Val	ACT Thr	TTA Leu 45	TTA Leu	ATG Met	GTT Val	144
CTT Leu	ACT Thr 50	CGT Ārg	ATT Ile	ATT Ile	GCA Ala	ATT Ile 55	GTG Val	TTT Phe	TTT Phe	AGA Arg	AGG Arg 60	CAA Gln	GAA Glu	CTC Leu	TTA Leu	192
CTT Leu 65	TGG Trp	CAG Gln	ATA Ile	GGT Gly	TGT Cys 70	GTT Val	CTG Leu	CTA Leu	GAG Glu	TAT Tyr 75	GGT Gly	AGT Ser	CCA Pro	AAA Lys	ATT Ile 80	240
AAA Lýs	TCC Ser	CTA Leu	GCA Ala	ATT Ile 85	AGC Ser	TTT Phe	TTA Leu	ACA Thr	GAA Glu 90	CTT Leu	TTT Phe	CAG Gln	CTT Leu	GGA Gly 95	GGA Gly	288
CTA Leu	CCA Pro	GCA Ala	CAA Gln 100	CCA Pro	GCT Ala	AGC Ser	ACT Thr	TTT Phe 105	TTC Phe	AGC Ser	TCA Ser	TTT Phe	TTG Leu 110	GAA Glu	TTA Leu	336
TTA Leu	AAA Lys	CAC His 115	CTT Leu	GTA Val	GAA Glu	ATG Met	GAT Asp 120	ACT Thr	GAC Asp	CAA Gln	TTG Leu	AAA Lys 125	CTC Leu	TAT Tyr	GAA Glu	384
GAG Glu	CCA Pro 130	TTA Leu	TCA Ser	AAG Lys	CTG Leu	ATA Ile 135	AAG Lys	ACA Thr	CTA Leu	TTT Phe	CCC Pro 140	TTT Phe	GAA Glu	GCA Ala	GAA Glu	432
GCT Ala 145	TAT Tyr	AGA Arg	AAT Asn	ATT Ile	GAA Glu 150	CCT Pro	GTC Val	TAT Tyr	TTA Leu	AAT Asn 155	ATG Met	CTG Leu	CTG Leu	GAA Glu	AAA Lys 160	480
CTC Leu	TGT Cys	GTC Val	ATG Met	TTT Phe 165	GAA Glu	GAC Asp	GGT Gly	GTG Val	CTC Leu 170	ATG Met	CGG Arg	CTT Leu	AAG Lys	TCT Ser 175	GAT Asp	528
TTG Leu	CTA Leu	AAA Lys	GCA Ala 180	GCT Ala	TTG Leu	TGC Cys	CAT His	TTA Leu 185	CTG Leu	CAG Gln	TAT Tyr	TTC Phe	CTT Leu 190	AAA Lys	TTT Phe	576
GTG Val	CCA Pro	GCT Ala 195	GGG Gly	TAT Tyr	GAA Glu	TCT Ser	GCT Ala 200	TTA Leu	CAA Gln	GTC Val	AGG Arg	AAG Lys 205	GTC Val	TAT Tyr	GTG Val	624
AGA Arg	AAT Asn 210	ATT Ile	TGT Cys	AAA Lys	GCT Ala	CTT Leu 215	TTG Leu	GAT Asp	GTG Val	CTT Leu	GGA Gly 220	ATT Ile	GAG Glu	GTA Val	GAT Asp	672

GCA GAG TAC TTG TTG GGC CCA CTT TAT GCA GCT TTG AAA ATG GAA AGT 220 225 230 225 236 227 227																	
Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Gln Glu Asn Leu 255 AGC AGT AAT AGT GAT GGA TGGA ATA TCA CCC AAA AGG CGT CGT CTC AGC TCG Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Arg Arg Leu Ser Ser 260 816 TCT CTA AAC CCT TCT AAA AGA GCA CCA AAA CAG ACT GAG GAA ATT AAA Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys 285 864 CCT GTG GAC ATG AAC CAA AAG AGC ACA ATA TTA TGG ACT GCA CTG AAA CAG HIS VALL Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln 295 912 AAA GCT GAA TCC CTT CAG ATT TCC CTT GAA TAC AGT GGC CTA AAG AAT LYS AS ALA GLU Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn 320 960 Lys Ala Glu Ser Leu Gln Ile Ser Leu Gly Ile Ala Val Val Leu Gln Leu Trp 325 310 CCT GTT ATT GAG ATG TA GAA GGA ATT GCT GTT GTC TTA CAA CTG ACT ACT ACT ACT ACT ACT ACT GTT CCT CAT CAA AAC ATG AAC TGC CGT ACT ACT ALA Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 340 1056 TCC AAG GAC TGT CAA CAT AAA TCC AAG AAC AAG AAC ATG ACC TGC GTA CT ACT ACT ACT ACT ACT ACT ACT ACT AC	Ala					Gly					Ala					Ser	720
Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Leu Ser Ser 265 TCT CTA AAC CCT TCT AAA AGA GCA CCA AAA CAG ACT GAG GAA ATT AAA Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys 275 CAT GTG GAC ATG AAC CAA AAG AGC ATA TTA TGG AGT GCA CTG AAA CAG 18 Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln 295 AAA GCT GAA TCC CTT CAG ATT TCC CTT GAA TAC AGT GGC CTA AAG AAT Lys Ala Glu Ser Leu Gln 11e Ser Leu Glu Tyr Ser Gly Leu Lys Asn 330 CCT GTT ATT GAG ATG TTA GAA GGA ATT GCT GTT GTC TTA CAA CTG ACT Lys Ala Glu Ser Leu Glu Glu Ile Ala Val Val Leu Glu Eu Trn 332 GCT CTG TATT GAG ATG TCT CAT TGT TCT CAT CAA AAC ATG AAC TGG CCT ACT ALA Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 345 TCC AAG GAC TGT CAA CAT AAA TCC AAG AAG AAA CCT TCT GTA GTG ATA Thr Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 375 AGT TTG TAT GAA TCT GTT CAC CAC AAA CTG GC CTT AAG ACT TGT TGT TGT TGT TGT TGT AGA GCT ATG ATG ATG ATG ATG ATG ATG ATG ATG AT					Glu					Gln					Asn		768
Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys 280 285 285 285 285 285 285 285 285 285 285	AGC Ser	AGT Ser	AAT Asn	Ser	GAT Asp	GGA Gly	ATA Ile	TCA Ser	Pro	AAA Lys	AGG Arg	CGT Arg	CGT Arg	Leu	AGC Ser	TCG Ser	816
# His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln 2995 AAA GCT GAA TCC CTT CAG ATT TCC CTT GAA TAC AGT GGC CTA AAG AAT 105	TCT Ser	CTA Leu	Asn	CCT Pro	TCT Ser	AAA Lys	AGA Arg	Ala	CCA Pro	AAA Lys	CAG Gln	ACT Thr	Glu	GAA Glu	ATT Ile	AAA Lys	864
Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Agn 320 CCT GTT ATT GAG ATG TTA GAA GGA ATT GCT GTT GTC TTA CAA CTG ACT Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr 325 GCT CTG TGT ACT GTT CAT TGT TCT CAT CAA AAC ATG AAC TGC CGT ACT ALa Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 340 TTC AAG GAC TGT CAA CAT AAA TCC AAG AAG AAA CCT TCT GTA GTA GTA ATG AAC Cys Gln His Lys Ser Lys Lys Pro Ser Val Val Ile 355 ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT TCT GTA GTG ATA 1104 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Pro Ser Val Val Ile 355 ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT AAG AGC TGT AGA TAC TTP TRP Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 375 AGT TTG TTA GAA TCT GTT CAG AAA CTG GAC CTG GAG GCA ACC ATT GAT Ser Leu Leu Glu Ser Val Gln Lys Leu Asp 395 AGG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT 400 AAG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT 405 TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TAT ATG CAA GTA AAC AGT 405 CCA TGG ATT TAT TAT CC CAT TCT GAT GAT GGC TGT TAAG ATG CTC TCA CTT Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420 CCA TGG ATT TAT TAT CC CAT TCT GAT GAT GGC TGT TAAG TTG ACC ACA ACC ACA ACC ACA ACC ACC ACC AC	CAT His	Val	GAC Asp	ATG Met	AAC Asn	CAA Gln	Lys	AGC Ser	ATA Ile	TTA Leu	TGG Trp	Ser	GCA Ala	CTG Leu	AAA Lys	CAG Gln	912
Pro Val Ile Glu Met Leu Glu Glu Gly Ile Ala Val Leu Gln Leu Thr 335 GCT CTG TGT ACT GTT CAT TGT TCT CAA AAC ATG AAC TGC CGT ACT ALA Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 340 TTC AAG GAC TGT CAA CAT AAA TCC AAG AAG AAG CCT TCT GTA GTG ATA ALA Leu Cys Asp Cys Gln His Lys Ser Lys Lys Pro Ser Val Val Ile 365 ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT AAG AGC TGT AGA AGC TTT TAC ACG AAA CTG GCC CYs Arg 377 ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT AAG AGC TGT AGA ACT ATA AGC TGT TAC ACA AAA ACT ACT ACT ACT ACT ACT AC	Lys	GCT Ala	GAA Glu	TCC Ser	CTT Leu	Gln	ATT Ile	TCC Ser	CTT Leu	GAA Glu	Tyr	AGT Ser	GGC Gly	CTA Leu	AAG Lys	Asn	960
Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 340 TTC AAG GAC TGT CAA CAT AAA TCC AAG AAG AAA CCT TCT GTA GTG ATA Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile 365 ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT AAG AGC TGT AGA TTT TTP Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 370 AGT TTG TTA GAA TCT GTT CAG AAA CTG GAC CTG GAG GCA ACC ATT GAT Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 390 AAG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT Leu Leu Glu Ser Val Gln Lys Leu Asp 395 TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TTA TGT GAA ATC ATC ATC ACT ACT ACT ACT ACT ACT A					Met					Ala					Leu		1008
Phe Lys Asp Cys Gln His Lys Ser Lys Lys Lys Pro Ser Val Val Ile 365 ACT TGG ATG TCA TTG GAT TTT TAC ACA AAA GTG CTT AAG AGC TGT AGA TTP Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 370 AGT TTG TTA GAA TCT GTT CAG AAA CTG GAC CTG GAG GCA ACC ATT GAT Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 385 AGG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser 415 TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TGT GGA ATG CTC TCA CTT Ser Phe Glu Asp Lie Glu Asp Leu Cys Gly Met Leu Ser Leu 420 CCA TGG ATT TAT TCC CAT TCT GAT GAT GAT GGC TGT TTA AAG TTG ACC ACA Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr 435 TCA CCA CAG GCA AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT 450 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG TTT CCA AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG TTT CAAG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACC ACA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC ACC TTG GCC CTG AGA AGA ATA TTC CTT GAG TGG AGA AC				Thr					His					Cys			1056
Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg 375 AGT TTG TTA GAA TCT GTT CAG AAA CTG GAC CTG GAG GCA ACC ATT GAT 1200 Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 400 AAG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT 405 TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TGT GGA ATG CTC TCA CTT 405 Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420 CCA TGG ATT TAT TCC CAT TCT GAT GAT GGC TGT TTA AAG TTG ACC ACA 1344 Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr 435 TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT 1392 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT TCA CTT 1392 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA 645 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488			Asp					Ser					Ser				1104
Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 400 AAG GTG GTG AAA ATT TAT GAT GCT TTG ATT TAT ATG CAA GTA AAC AGT Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser 405 TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TGT GGA ATG CTC TCA CTT Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420 CCA TGG ATT TAT TCC CAT TCT GAT GAT GGC TGT TTA AAG TTG ACC ACA Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr 435 TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 465 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488		\mathtt{Trp}					Phe					Leu					1152
Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser 415 TCA TTT GAA GAT CAT ATC CTG GAA GAT TTA TGT GGA ATG CTC TCA CTT Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420 CCA TGG ATT TAT TCC CAT TCT GAT GAT GGC TGT TTA AAG TTG ACC ACA Pro Trp Ile Tyr Ser His Ser Asp Gly Cys Leu Lys Leu Thr Thr 435 TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 480 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG AGG ATG TTP Ala Leu Cu Thr Ala Val Tyr Asn Trp Ala Leu	Ser					Val					Leu					Asp	1200
Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu 420 CCA TGG ATT TAT TCC CAT TCT GAT GAT GGC TGT TTA AAG TTG ACC ACA Pro Trp Ile Tyr Ser His Ser Asp Gly Cys Leu Lys Leu Thr Thr 435 TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 470 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG 1488 Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu					Ile					Ile					Asn		1248
Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr 435 TTT GCC GCT AAT CTT CTA ACA TTA AGC TGT AGG ATT TCA GAT AGC TAT Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 470 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG ACT AGA AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG ACT AGA AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG ACT AGA AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG ACT AGA AGA AGA AGA AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA AGA AGA AGA AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA ACT AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA ACT AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG ACC CTG ACT AGA ACT				Asp					Asp					Leu			1296
Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 450 TCA CCA CAG GCA CAA TCA CGA TGT GTG TTT CTT CTG ACT CTG TTT CCA Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 465 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu			Ile					Asp					Lys				1344
Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 465 470 480 AGA AGA ATA TTC CTT GAG TGG AGA ACA GCA GTT TAC AAC TGG GCC CTG Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu	TTT Phe	Ala	GCT Ala	AAT Asn	CTT Leu	CTA Leu	Thr	TTA Leu	AGC Ser	TGT Cys	AGG Arg	Ile	TCA Ser	GAT Asp	AGC Ser	TAT Tyr	1392
Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu	Ser	CCA Pro	CAG Gln	GCA Ala	CAA Gln	Ser	CGA Arg	TGT Cys	GTG Val	TTT Phe	Leu	CTG Leu	ACT Thr	CTG Leu	TTT Phe	Pro	1440
					Leu					Ala					Ala		1488

-167-

												AGT Ser			TTT Phe		1536
												CCC Pro 525					1584
ATA Ile	GAT Asp 530	AAA Lys	GTC Val	AAA Lys	GAT Asp	GAT Asp 535	TCT	GAC Asp	ATT Ile	GTC Val	AAG Lys 540	AAA Lys	GAA Glu	TTT Phe	GCT Ala		1632
TCT Ser 545	ATA Ile	CTT Leu	GGT Gly	CAA Gln	CTT Leu 550	GTC Val	TGT Cys	ACT Thr	CTT	CAC His 555	GGC Gly	ATG Met	TTT Phe	TAT Tyr	CTG Leu 560		1680
ACA Thr	AGT Ser	TCT Ser	TTA Leu	ACA Thr 565	GAA Glu	CCT Pro	TTC Phe	TCT Ser	GAA Glu 570	CAC His	GGA Gly	CAT His	GTG Val	GAC Asp 575	CTC Leu		1728
TTC Phe	TGT Cys	AGG Arg	AAC Asn 580	TTG Leu	AAA Lys	GCC Ala	ACT Thr	TCT Ser 585	CAA Gln	CAT His	GAA Glu	TGT Cys	TCA Ser 590	TCT Ser	TCT Ser		1776
												CTA Leu 605					1824
												AAT Asn					1872
CTT Leu 625	TGT Cys	AAG Lys	CAT His	CTT Leu	GAT Asp 630	TTT Phe	AGA Arg	GAA Glu	GAT Asp	GAA Glu 635	ACA Thr	GAT Asp	GTA Val	AAA Lys	GCA Ala 640		1920
GTT Val	CTT Leu	GGA Gly	ACT Thr	TTA Leu 645	TTA Leu	AAT Asn	TTA Leu	ATG Met	GAA Glu 650	GAT Asp	CCA Pro	GAC Asp	AAA Lys	GAT Asp 655	GTT Val		1968
AGA Arg	GTG Val	GCT Ala	TTT Phe 660	AGT Ser	GGA Gly	AAT Asn	ATC Ile	AAG Lys 665	CAC His	ATA Ile	TTG Leu	GAA Glu	TCC Ser 670	TTG Leu	GAC Asp		2016
TCT Ser	GAA Glu	GAT Asp 675	GGA Gly	TTT Phe	ATA Ile	AAG Lys	GAG Glu 680	CTT Leu	TTT Phe	GTC Val	TTA Leu	AGA Arg 685	ATG Met	AAG Lys	GAA Glu		2064
GCA Ala	TAT Tyr 690	ACA Thr	CAT His	GCC Ala	CAA Gln	ATA Ile 695	TCA Ser	AGA Arg	AAT Asn	AAT Asn	GAG Glu 700	CTG Leu	AAG Lys	GAT Asp	ACC Thr		2112
TTG Leu 705	ATT Ile	CTT Leu	ACA Thr	ACA Thr	GGG Gly 710	GAT Asp	ATT Ile	GGA Gly	AGG Arg	GCC Ala 715	GCA Ala	AAA Lys	GGA Gly	GAT Asp	TTG Leu 720		2160
GTA Val	CCA Pro	TTT Phe	GCA Ala	CTC Leu 725	TTA Leu	CAC His	TTA Leu	TTG Leu	CAT His 730	TGT Cys	TTG Leu	TTA Leu	TCC Ser	AAG Lys 735	TCA Ser	:	2208
GCA Ala	TCT Ser	GTC Val	TCT Ser 740	GGA Gly	GCA Ala	GCA Ala	TAC Tyr	ACA Thr 745	GAA Glu	ATT Ile	AGA Arg	GCT Ala	CTG Leu 750	GTT Val	GCA Ala	:	2256
GCT Ala	AAA Lys	AGT Ser 755	GTT Val	AAA Lys	CTG Leu	CAA Gln	AGT Ser 760	TTT Phe	TTC Phe	AGC Ser	CAG Gln	TAT Tyr 765	AAG Lys	AAA Lys	CCC Pro	:	2304

ATC Ile	TGT Cys 770	CAG Gln	TTT Phe	TTG Leu	GTA Val	GAA Glu 775	TCC Ser	CTT Leu	CAC His	TCT Ser	AGT Ser 780	CAG Gln	ATG Met	ACA Thr	GCA Ala	2352
CTT Leu 785	CCG Pro	AAT Asn	ACT Thr	CCA Pro	TGC Cys 790	CAG Gln	AAT Asn	GCT Ala	GAC Asp	GTG Val 795	CGA Arg	AAA Lys	CAA Gln	GAT Asp	GTG Val 800	2400
GCT Ala	CAC His	CAG Gln	AGA Arg	GAA Glu 805	ATG Met	GCT Ala	TTA Leu	AAT Asn	ACG Thr 810	TTG Leu	TCT Ser	GAA Glu	ATT Ile	GCC Ala 815	AAC Asn	2448
GTT Val	TTC Phe	GAC Asp	TTT Phe 820	CCT Pro	GAT Asp	CTT Leu	AAT Asn	CGT Arg 825	TTT Phe	CTT Leu	ACT	AGG Arg	ACA Thr 830	TTA Leu	CAA Gln	2496
GTT Val	.CTA Leu	CTA Leu 835	CCT Pro	GAT Asp	CTT Leu	GCT Ala	GCC Ala 840	AAA Lys	GCA Ala	AGC Ser	CCT Pro	GCA Ala 845	GCT Ala	TCT Ser	GCT Ala	2544
CTC Leu	ATT Ile 850	CGA Arg	ACT Thr	TTA Leu	GGA Gly	AAA Lys 855	CAA Gln	TTA Leu	AAT Asn	GTC Val	AAT Asn 860	CGT Arg	AGA Arg	GAG Glu	ATT Ile	2592
TTA Leu 865	ATA Ile	AAC Asn	AAC Asn	TTC Phe	AAA Lys 870	TAT Tyr	ATT Ile	TTT Phe	TCT Ser	CAT His 875	TTG Leu	GTC Val	TGT Cys	TCT Ser	TGT Cys 880	2640
TCC Ser	AAA Lys	GAT Asp	GAA Glu	TTA Leu 885	GAA Glu	CGT Arg	GCC Ala	CTT Leu	CAT His 890	TAT Tyr	CTG Leu	AAG Lys	AAT Asn	GAA Glu 895	ACA Thr	2688
Glu	Ile	Glu	Leu 900	Gly	AGC Ser	Leu	Leu	Arg 905	Gln	Asp	Phe	Gln	Gly 910	Leu	His	2736
AAT Asn	GAA Glu	TTA Leu 915	TTG Leu	CTG Leu	CGT Arg	ATT Ile	GGA Gly 920	GAA Glu	CAC His	TAT Tyr	CAA Gln	CAG Gln 925	GTT Val	TTT Phe	AAT Asn	2784
GGT Gly	TTG Leu 930	TCA Ser	ATA Ile	CTT Leu	GCC Ala	TCA Ser 935	TTT Phe	GCA Ala	TCC Ser	AGT Ser	GAT Asp 940	GAT Asp	CCA Pro	TAT Tyr	CAG Gln	2832
GGC Gly 945	CCG Pro	AGA Arg	GAT Asp	ATC Ile	ATA Ile 950	TCA Ser	CCT Pro	GAA Glu	CTG Leu	ATG Met 955	GCT Ala	GAT Asp	TAT Tyr	TTA Leu	CAA Gln 960	2880
CCC Pro	AAA Lys	TTG Leu	TTG Leu	GGC Gly 965	ATT	TTG Leu	GCT Ala	TTT Phe	TTT Phe 970	AAC Asn	ATG Met	CAG Gln	TTA Leu	CTG Leu 975	AGC Ser	2928
TCT Ser	AGT Ser	GTT Val	GGC Gly 980	ATT Ile	GAA Glu	GAT Asp	AAG Lys	AAA Lys 985	ATG Met	GCC Ala	TTG Leu	AAC Asn	AGT Ser 990	TTG Leu	ATG Met	2976
TCT Ser	TTG Leu	ATG Met 995	AAG Lys	TTA Leu	ATG Met	GGA Gly	CCC Pro 1000	Lys	CAT His	GTC Val	AGT Ser	TCT Ser 1005	Val	AGG Arg	GTG Val	3024
AAG Lys	ATG Met 1010	Met	ACC Thr	ACA Thr	CTG Leu	AGA Arg 1015	Thr	GGC Gly	CTT Leu	CGA Arg	TTC Phe 1020	Lys	GAT Asp	GAT Asp	TTT Phe	3072
CCT Pro 1025	Glu	TTG Leu	TGT Cys	TGC Cys	AGA Arg 1030	Ala	TGG Trp	GAC Asp	TGC Cys	TTT Phe 1035	Val	CGC Arg	TGC Cys	CTG Leu	GAT Asp 1040	3120

					Ser					Val			GCT Ala		Leu	3168
				Ile					Thr				TTC Phe 107	His		3216
			Glu					Val					CAT His 5			3264
TAT Tyr	TTT Phe 109	Leu	CCT Pro	GAT Asp	CAT His	CCA Pro 1099	Glu	TTA Leu	AAA Lys	AAG Lys	ATA Ile 1100	Lys	GCC Ala	GTT Val	CTC Leu	3312
	Glu					Thr					Asp		CAG Gln			3360
CTT Leu	CAG Gln	CTC Leu	TCT Ser	ATG Met 1125	Lys	GCC Ala	ATT Ile	CAA Gln	CAT His 1130	Glu	AAT Asn	GTC Val	GAT Asp	GTT Val 1135	Arg	3408
ATT Ile	CAT His	GCT Ala	CTT Leu 1140	Thr	AGC Ser	TTG Leu	AAG Lys	GAA Glu 1145	Thr	TTG Leu	TAT Tyr	AAA Lys	AAT Asn 1150	Gln	GAA Glu	3456
AAA Lys	CTG Leu	ATA Ile 1155	Lys	TAT Tyr	GCA Ala	ACA Thr	GAC Asp 1160	Ser	GAA Glu	ACA Thr	GTA Val	GAA Glu 116	CCT Pro	ATT Ile	ATC Ile	3504
TCA Ser	CAG Gln 1170	Leu	GTG Val	ACA Thr	GTG Val	CTT Leu 1175	Leu	AAA Lys	GGT Gly	TGC Cys	CAA Gln 1180	Asp	GCA Ala	AAC Asn	TCT Ser	3552
CAA Gln 1189	Ala	CGG Arg	TTG Leu	CTC Leu	TGT Cys 1190	Gly	GAA Glu	TGT Cys	TTA Leu	GGG Gly 1195	Glu	TTG Leu	GGG Gly	GCG Ala	ATA Ile 1200	3600
GAT Asp	CCA Pro	GGT Gly	CGA Arg	TTA Leu 1205	Asp	TTC Phe	TCA Ser	ACA Thr	ACT Thr 1210	Glu	ACT Thr	CAA Gln	GGA Gly	AAA Lys 1215	Asp	3648
				Thr					Ser				TAT Tyr 1230	Gly		3696
TTG Leu	ATG Met	GAG Glu 1235	Leu	ACA Thr	AGA Arg	GCT Ala	TAC Tyr 1240	Leu	GCG Ala	TAT Tyr	GCT Ala	GAT Asp 1245	AAT Asn	AGC Ser	CGA Arg	3744
GCT Ala	CAA Gln 1250	Asp	TCA Ser	GCT Ala	GCC Ala	TAT Tyr 1255	Ala	ATT Ile	CAG Gln	GAG Glu	TTG Leu 1260	Leu	TCT Ser	ATT Ile	TAT Tyr	3792
GAC Asp 1265	Cys	AGA Arg	GAG Glu	ATG Met	GAG Glu 1270	Thr	AAC Asn	GGC Gly	CCA Pro	GGT Gly 1275	His	CAA Gln	TTG Leu	TGG Trp	AGG Arg 1280	3840
AGA Arg	TTT Phe	CCT Pro	GAG Glu	CAT His 1285	Val	CGG Arg	GAA Glu	ATA Ile	CTA Leu 1290	Glu	CCT Pro	CAT His	CTA Leu	AAT Asn 1295	Thr	3888
AGA Arg	TAC Tyr	AAG Lys	AGT Ser 1300	Ser	CAG Gln	AAG Lys	TCA Ser	ACC Thr 1305	Asp	TGG Trp	TCT Ser	GGA Gly	GTA Val 1310	Lys	AAG Lys	3936

CCA ATT TAC TTA AGT AAA TTG GGT AGT AAC TTT GCA GAA TGG TCA GCA PRO ILE TRY Leu Ser Lys Leu Gly Ser Aen Phe Ala Glu TTP Ser Ala 1315 1315 1320 4032 1325 1340 1325 1340 1325 1340 1325 1340 1340 1355 1340 1355 1340 1355 1340 1355 1340 1355 1340 1355 1340 1355 1340 1355 1340 1355 1340 1355	•																
Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1335 1340 1340 1340 1340 1340 1340 1345 1340 1350 1360 1345 1350 1360 1345 1350 1360 1345 1350 1360 1345 1350 1365 1366 1365 1366 1365 1366 1365 1366 1370	CCA Pro	ATT Ile	Tyr	Leu	AGT Ser	AAA Lys	TTG Leu	Gly	Ser	AAC Asn	TTT Phe	GCA Ala	Glu	Trp	TCA Ser	GCA Ala	3984
Lys 1le Phe Thr Cys Cys Ser 1le Met Met Lys His Asp Phe Lys Val 1345 ACC ATC TAT CTT CCA CAT ATT CTG GTG TAT GTG GTG TGT TTR TIle Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1375 AAT CAA GAA GAT CAG CAG GAG GTT TAT GCA GAA ATT ATG GCA GTT CTA Ash Gln Glu Asp Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1390 AAG CAT GAC GAT CAC ATA AAT ACC CAA GAC ATT GCA TCT GAT CAT Lys His Asp Asp Asp Gln His Thr Ile Ash Thr Gln Asp Ile Ala Ser Asp Asp 1405 CTG TGT CAA CTC AGT ACC ATA AAT ACC CAT CTT GAC CAT CTC CTC CYs Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu Lat Lat Leu Lys Gln Leu Ser Thr Gln Thr Ala Asp His Leu Lat Leu Lys Ala Glu Lys Cys Lat Lat Lys Leu Lys Leu Lys Lat Lys Lat Lat Lat Lys Lat Lat Lat Lys Lat Lat Lat Lys Lat		Trp	Ala				Ile	Thr				His	Asp				4032
Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1375 AAT CAA GAA GAT CAG CAG GAG GTT TAT GCA GAA ATT ATG GCA GTT CTA Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 AAG CAT GAC GAT CAG CAT ACC ATA AAT ACC CAA GAC ATT GCA TCT GAT Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395 CTG TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG CTT GAC CAT CTC Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 ACA CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA GCT GAG AAA TGT Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTT ACT Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 CCC CAG GAT ACT CTG GCA GTA GCT TCC TCT TC GC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT AFG GAA CAC Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1480 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT AFG GAA CCA GAA CCA TTT GAA TCA CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT AFG GAA CCA GAA CCT GAA GAA CCA TTT GAA TCA TTT TATA CAG GAA AAG AAG CAA AAT AFG GAA CCT GAA GAA CAC TTT GAA TCA TTT TATA CAG GAA TTA AGA AAG GCA TAC ACA 1495 ATT CAG GAA CAT CTT GAA TTT TATA CAG GAA ATG AGA CAA AAT TAG GAA CCT GAA GAG CAA CTT GAA CCT GAA CCT GAA CCT GAA CAC TTT GAA TCA TT TAG CAG GAA CCT GAA CCA GAA CCT GAA CCA GAA CCT GAA CCT GAA CCA GAA CCT GAA CCT GAA CCA GAA CCT GAA CCA GAA CCT GAA CC	Lys	Ile				Cys	Ser				Lys	His				Val	4080
Ash Gin Glu Asp Gin Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 AAG CAT GAC GAT CAG CAT ACC ATA AAT ACC CAA GAC ATT GCA TCT GAT Lys His Asp Asp Gin His Thr Ile Ash Thr Gin Asp Ile Ala Ser Asp 1395 CTG TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG CTT GAC CAT CTC Leu Cys Gin Leu Ser Thr Gin Thr Val Phe Ser Met Leu Asp His Leu 1410 ACA CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA GCT GAG AAA TGT THR Gin Trp Ala Arg His Lys Phe Gin Ala Leu Lys Ala Glu Lys Cys 1425 CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT Pro His Ser Lys Ser Ash Arg Ash Lys Val Asp Ser Met Val Ser Thr 1445 GTG GAT TAT GAA GAC TAT CAG AGT GTA TCC TT TT CTG GAC CTC ATA CAT ASP TY Glu Asp Tyr Thr 1460 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA ATG TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA TAC TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe 11e Thr Glu Lys Lys Gln Ash 1495 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT ARg Ala Val Met His Phe Glu Ser Phe 11e Thr Glu Lys Lys Gln Ash 1495 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT 11e Gln Glu His Leu Gly Phe Leu Gln Lys Ala Ala Met His 1515 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala Tle Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA GCC TTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAA CAG GCT ATT CAG CTA GAA CCA GAA CCA GLU Pro Asp Gly Val Ala Gly Val Ser Ala Tle Glu Glu Pro Asp Gly Val Ala Gly Val Ser Ala Tle Glu Glu Pro Asp Gly Val Ala Gly Val Ser Ala Tle Glu Glu Pro Asp Gly 1540 GAT GCC ACT GCT TGT TAT GAA CAG GCT ATT CAG CTA GAA CCA GAC CAG ASP Ala Thr Ala Cys Tyr Asp Arg Ala Ile Glu Leu Glu Pro Asp Glu 1540					Leu	Pro				Val	Tyr				Gly	Cys	4128
Lys His Asp Asp Gln His Thr 11e Asn Thr Gln Asp 11e Ala Ser Asp 1395 CTG TGT CAA CTC AGT ACA CAG ACT GTG TTC TCC ATG CTT GAC CAT CTC Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 ACA CAG TGG GCA AGG CAC AAA TTT CAG GCA CTG AAA GCT GAG AAA TGT Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 CCA CAC AGC AAA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu 11e 1460 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CCG AGT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT AFG Ala Val Met His Phe Glu Ser Phe 11e Thr Glu Lys Lys Gln Asn 1495 CCA GCT GTA ATG CAC TTT GAA TCA TTT TAT ACA GAA AAG CAA TAC ACA Pro Gln Ser Cac TTT TAT ACA GAA AAG AAG CAA AAT AFG Ala Val Met His Phe Glu Ser Phe 11e Thr Glu Lys Lys Gln Asn 1495 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT AFG Ala Val Met His Phe Glu Ser Phe 11e Thr Glu Lys Lys Gln Asn 1495 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT T1E GIn Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1500 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala 11e Arg Lys Ala Glu Pro 1535 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GCC TTG CTG AGG Ser Leu Lys Glu Gln file Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1530 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG ASP Ala Thr Ala Cys Tyr Asp Arg Ala 11e Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AGT CAT GTA GTA GGT CTT GGT CAG T16e Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gin His Leu Gly Val Val Lys Ser Met Leu Gly Leu Gly Gin				Asp	Gln				Tyr	Ala				Ala	Val		4176
Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 ACA CAG TGG GCA AGG CAC AGA TTT CAG GCA CTG AGA GCT GAG AAA TGT Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 CCA CAC AGC AGA TCA AGC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1470 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1485 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT ARG Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1495 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT 1490 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT 1520 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA ACA GCA His Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1510 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA GLU Pro Asp Gly Val Ala Gly Val Ser Ala 11e Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA CAT GAA AGC CTT GGC TTG CTG AGG CAG GAC CAG GAA CCA GAC CAT GCA GAC CAG GAA CCA GAC CAT GCA GCA GAC CAT GCA GAC CAT GCA GCA GAC CAT GCA GCA GAC CAT GCA GCA GAC GAC GAC GAC GAC GAC GAC GAC			Asp	Asp				Ile	Asn				Ile	Ala			4224
Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1435 CCA CAC AGC AGA TCA AAC AGA AAT AAG GTA GAC TCA ATG GTA TCT ACT 1446 Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1455 GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA 1416 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA 1416 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA 1464 Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT 1490 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT 1500 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA GIU Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TGG AGG CAG AAT 1520 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG GAA AGC CCA GAC CAG GAT CAT GCT CTT GAA GAA CAT CTT GAA AGA AGC CTT GCT GCT ATG CAT 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG CTG CTG CTG CTG AGG CTG CTG CTG CTG CTG AGG CTG CTG CTG CTG CTG CTG CTG CTG CTG C		Cys	Gln				Gln	Thr				Met	Leu				4272
Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1455 GTG GAT TAT GAA GAC TAT CAG AGT GTA ACC CGT TTT CTA GAC CTC ATA AAA GAC ASP Leu Ile 1460 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1510 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA GRU Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG ASP Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1550 ATC ATT CAT TAC CAT GGT GTA GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG CAG ASP ATT AGA AND GLA CAT GAT GAS AGG GCA CAG CAG ATT AGA AND CAG GCT TGT TAT GAC AGG GCT ATT CAG CTA GAC CAG ASP ATT AGA AND CAG GCT GCT TGT TAT GAC AGG GCT AGG CTA GAC CAG ASP ATT AGA AGG CCTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG ASP ATT AGA ALA CTA GAT CAT GAT TAT AGA AGG CTA GAT CAG CAG CAG ASP ATT AGA ALA CTA GAT CAT GAT CAG CAG CAG ASP ATT AGA ALA CTA GAT TAT AGA AGG CTA GAT CAG CAG CAG ASP ATT AGA ALA CTA GAT TAT AGA AGG CTA GAT CAG CAG CAG ASP ATT AGA ALA CTA GAT TAT AGA AGG CTA GAT CAG CAG CAG ACC AGA CAG ACC AGG ACC AGG ACC AGG CAG ACC AGG ACC AGG ACC AGG ACC AGA ACC AGG	Thr	Gln				His	Lys				Leu	Lys				Cys	4320
Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1470 CCC CAG GAT ACT CTG GCA GTA GCT TCC TTT CGC TCC AAA GCA TAC ACA Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 ATT CAG GAA CAT CTT GGA TTT TA CAG AAA TTG TAT GCT GCT ATG CAT Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1535 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 4752					Ser	Asn				Val	Asp				Ser	Thr	4368
Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 CGA GCT GTA ATG CAC TTT GAA TCA TTT ATT ACA GAA AAG AAG CAA AAT Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTG GGT CAG 1752 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG 1752 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG 1752 Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1752				Glu	Asp				Val	Thr				Asp	Leu		4416
ATT CAG GAA CAT CTT GGA TTT TTA CAG AAA TTG TAT GCT GCT ATG CAT Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 4560 4560 4560 4560 4560 4764 4762			Asp	Thr				Ala	Ser				Lys	Ala			4464
The Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 GAA CCT GAT GGA GTG GCC GGA GTC AGT GCA ATT AGA AAG GCA GAA CCA 4608 Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG 1535 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG 4656 Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG 4704 Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG 1752 Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln		Ala	Val				Glu	Ser				Glu	Lys				4512
Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 TCT CTA AAA GAA CAG ATC CTT GAA CAT GAA AGC CTT GGC TTG CTG AGG Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln	Ile	Gln				Gly	Phe				Leu	Tyr				His	4560
Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 GAT GCC ACT GCT TGT TAT GAC AGG GCT ATT CAG CTA GAA CCA GAC CAG Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln	GAA Glu	CCT Pro	GAT Asp	GGA Gly	Val	Ala	GGA Gly	GTC Val	AGT Ser	Ala	Ile	AGA Arg	AAG Lys	GCA Ala	Glu	Pro	4608
Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 ATC ATT CAT TAC CAT GGT GTA GTA AAG TCC ATG TTA GGT CTT GGT CAG Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln				Glu	Gln				His	Glu				Leu	Leu		4656
Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln			Thr	Ala				Arg	Ala				Glu	Pro			4704
		Ile	His				Val	Val				Leu	Gly				4752

	Ser					Gln			GGA Gly		His				TCC Ser 1600	4800
					Leu				AGA Arg 161	Val						4848
				Asp					Tyr					Gly	AAA Lys	4896
			Trp					Gly	CAG Gln				Ser		AAA Lys	4944
		Asp					Tyr					Leu			GCA Ala	4992
	Gln					Ser			AGC Ser		Glu					5040
CAA Gln	CGA Arg	GGA Gly	TAT Tyr	GAA Glu 1689	Tyr	ATT Ile	GTG Val	AGA Arg	TTG Leu 1690	His	ATG Met	TTA Leu	TGT Cys	GAG Glu 169	Leu	5088
				Lys					CAT His					Ser		5136
			Ser					Ala	CGA Arg				Thr			5184
		Arg					Ile		GCT Ala			Arg				5232
AGC Ser 1745	Leu	AAC Asn	AAA Lys	AGA Arg	CCA Pro 1750	Asp	TAC Tyr	AAT Asn	GAA Glu	ATG Met 1755	Val	GGA Gly	GAA Glu	TGC Cys	TGG Trp 1760	5280
					Val				GCT Ala 1770	Gly					Ala	5328
TAC Tyr	AAT Asn	GCT Ala	CTC Leu 1780	Leu	AAT Asn	GCA Ala	GGG Gly	GAA Glu 1789	TCA Ser	CGA Arg	CTC Leu	GCT Ala	GAA Glu 1790	Leu	TAC Tyr	5376
GTG Val	GAA Glu	AGG Arg 1799	Ala	AAG Lys	TGG Trp	CTC Leu	TGG Trp 1800	Ser	AAG Lys	GGT Gly	GAT Asp	GTT Val 1805	His	CAG Gln	GCA Ala	5424
CTA Leu	ATT Ile 1810	Val	CTT Leu	CAA Gln	AAA Lys	GGT Gly 1815	Val	GAA Glu	TTA Leu	TGT Cys	TTT Phe 1820	Pro	GAA Glu	AAT Asn	GAA Glu	5472
ACC Thr 1825	Pro	CCT Pro	GAG Glu	GGT Gly	AAG Lys 1830	Asn	ATG Met	TTA Leu	ATC Ile	CAT His 1835	Gly	CGA Arg	GCT Ala	ATG Met	CTA Leu 1840	5520
CTA Leu	GTG Val	GGC Gly	CGA Arg	TTT Phe 1845	Met	GAA Glu	GAA Glu	ACA Thr	GCT Ala 1850	Asn	TTT Phe	GAA Glu	AGC Ser	AAT Asn 1855	Ala	5568

ATT ATG AAA Ile Met Lys	AAA TAT AAG Lys Tyr Lys 1860	Asp Val Th	C GCG TGC C r Ala Cys Le 65	TG CCA GAA TGG eu Pro Glu Trp 1870	GAG 5616 Glu
GAT GGG CAT Asp Gly His 1875	Phe Tyr Leu	GCC AAG TA Ala Lys Ty 1880	C TAT GAC AM	AA TTG ATG CCC ys Leu Met Pro 1885	ATG 5664 Met
GTC ACA GAC Val Thr Asp 1890	AAC AAA ATG Asn Lys Met	GAA AAG CA Glu Lys Gl 1895	n Gly Asp Le	TC ATC CGG TAT eu Ile Arg Tyr 900	ATA 5712 Ile
GTT CTT CAT Val Leu His 1905	TTT GGC AGA Phe Gly Arg 191	Ser Leu Gl	A TAT GGA AA n Tyr Gly As 1915	AT CAG TTC ATA sn Gln Phe Ile	TAT 5760 Tyr 1920
CAG TCA ATG Gln Ser Met	CCA CGA ATG Pro Arg Met 1925	TTA ACT CT Leu Thr Le	A TGG CTT GA u Trp Leu As 1930	AT TAT GGT ACA sp Tyr Gly Thr 193!	Lys
GCA TAT GAA Ala Tyr Glu	TGG GAA AAA Trp Glu Lys 1940	GCT GGC CG Ala Gly Ar 19	g Ser Asp Ar	GT GTA CAA ATG rg Val Gln Met 1950	AGG 5856 Arg
AAT GAT TTG Asn Asp Leu 1955	Gly Lys Ile	AAC AAG GT Asn Lys Va 1960	T ATC ACA GA l lle Thr Gl	AG CAT ACA AAC Lu His Thr Asn 1965	TAT 5904 Tyr
TTA GCT CCA Leu Ala Pro 1970	TAT CAA TTT Tyr Gln Phe	TTG ACT GC Leu Thr Al 1975	a Phe Ser Gl	AA TTG ATC TCT In Leu Ile Ser 980	CGA 5952 Arg
ATT TGT CAT Ile Cys His 1985	TCT CAC GAT Ser His Asp 199	Glu Val Ph	T GTT GTC TI e Val Val Le 1995	TG ATG GAA ATA eu Met Glu Ile	ATA 6000 Ile 2000
GCC AAA GTA Ala Lys Val	TTT CTA GCC Phe Leu Ala 2005	TAT CCT CA. Tyr Pro Gl	A CAA GCA AT n Gln Ala Me 2010	TG TGG ATG ATG et Trp Met Met 2015	Thr
GCT GTG TCA Ala Val Ser	AAG TCA TCT Lys Ser Ser 2020	TAT CCC ATO Tyr Pro Me 20	t Arg Val As	AC AGA TGC AAG sn Arg Cys Lys 2030	GAA 6096 Glu
ATC CTC AAT Ile Leu Asn 2035	Lys Ala Ile	CAT ATG AA His Met Ly: 2040	A AAA TCC TI s Lys Ser Le	TA GAG AAG TTT eu Glu Lys Phe 2045	GTT 6144 Val
GGA GAT GCA Gly Asp Ala 2050	ACT CGC CTA Thr Arg Leu	ACA GAT AAG Thr Asp Lys 2055	s Leu Leu Gl	AA TTG TGC AAT Lu Leu Cys Asn 060	AAA 6192 Lys
CCG GTT GAT Pro Val Asp 2065	GGA AGT AGT Gly Ser Ser 2070	Ser Thr Le	A AGC ATG AG u Ser Met Se 2075	GC ACT CAT TTT er Thr His Phe	AAA 6240 Lys 2080
ATG CTT AAA Met Leu Lys	AAG CTG GTA Lys Leu Val 2085	GAA GAA GC Glu Glu Ala	A ACA TTT AG a Thr Phe Se 2090	GT GAA ATC CTC er Glu Ile Leu 2095	Ile
Pro Leu Gln	TCA GTC ATG Ser Val Met 2100	ATA CCT ACT Ile Pro The 210	r Leu Pro Se	CA ATT CTG GGT er Ile Leu Gly 2110	ACC 6336 Thr
CAT GCT AAC His Ala Asn 2115	His Ala Ser	CAT GAA CCI His Glu Pro 2120	A TTT CCT GG o Phe Pro Gl	GA CAT TGG GCC Ly His Trp Ala 2125	TAT 6384 Tyr

ATT Ile	GCA Ala 213	Gly	TTT Phe	GAT Asp	GAT Asp	ATG Met 213	Val	GAA Glu	ATT	CTT Leu	GCT Ala 214	Ser	CTT	CAG Gln	AAA Lys	6432
CCA Pro 214	Lys	AAG Lys	ATT Ile	TCT Ser	TTA Leu 215	Lys	GGC Gly	TCA Ser	GAT Asp	GGA Gly 215	Lys	TTC Phe	TAC Tyr	ATC	ATG Met 2160	6480
ATG Met	TGT Cys	AAG Lys	CCA Pro	AAA Lys 216	Asp	GAC Asp	CTG Leu	AGA Arg	AAG Lys 217	Asp	TGT Cys	AGA Arg	CTA Leu	Met 217	GAA Glu 5	6528
TTC Phe	AAT Asn	TCC Ser	TTG Leu 218	Ile	AAT Asn	AAG Lys	TGC Cys	TTA Leu 218	Arg	AAA Lys	GAT Asp	GCA Ala	GAG Glu 219	Ser	CGT Arg	6576
AGA Arg	AGA Arg	GAA Glu 219	Leu	CAT His	ATT Ile	CGA Arg	ACA Thr 220	Tyr	GCA Ala	GTT Val	ATT Ile	CCA Pro 220	Leu	AAT Asn	GAT Asp	6624
GAA Glu	TGT Cys 221	GGG Gly 0	ATT Ile	ATT Ile	GAA Glu	TGG Trp 2219	Val	AAC Asn	AAC Asn	ACT Thr	GCT Ala 2220	Gly	TTG Leu	AGA Arg	CCT Pro	6672
ATT Ile 222	Leu	ACC Thr	AAA Lys	CTA Leu	TAT Tyr 2230	Lys	GAA Glu	AAG Lys	GGA Gly	GTG Val 2235	Tyr	ATG Met	ACA Thr	GGA Gly	AAA Lys 2240	6720
GAA Glu	CTT Leu	CGC Arg	CAG Gln	TGT Cys 224	Met	CTA Leu	CCA Pro	AAG Lys	TCA Ser 225	Ala	GCT Ala	TTA Leu	TCT Ser	GAA Glu 225	Lys	6768
CTC Leu	AAA Lys	GTA Val	TTC Phe 2260	Arg	GAA Glu	TTT Phe	CTC Leu	CTG Leu 2265	Pro	AGG Arg	CAT His	CCT Pro	CCT Pro 2270	Ile	TTT Phe	6816
CAT His	GAG Glu	TGG Trp 2275	Phe	CTG Leu	AGA Arg	ACA Thr	TTC Phe 2280	Pro	GAT Asp	CCT Pro	ACA Thr	TCA Ser 2285	Trp	TAC Tyr	AGT Ser	6864
AGT Ser	AGA Arg 229(TCA Ser	GCT Ala	TAC Tyr	TGC Cys	CGT Arg 2295	Ser	ACT Thr	GCA Ala	GTA Val	ATG Met 2300	Ser	ATG Met	GTT Val	GGT Gly	6912
TAT Tyr 2305	Ile	CTG Leu	GGG Gly	CTT Leu	GGA Gly 2310	Asp	CGT Arg	CAT His	GGT Gly	GAA Glu 2315	Asn	ATT Ile	CTC Leu	TTT Phe	GAT Asp 2320	6960
TCT Ser	TTG Leu	ACT Thr	GGT Gly	GAA Glu 2325	Cys	GTA Val	CAT His	GTA Val	GAT Asp 2330	Phe	AAT Asn	TGT Cys	CTT Leu	TTC Phe 2335	Asn	7008
AAG Lys	GGA Gly	GAA Glu	ACC Thr 2340	Phe	GAA Glu	GTT Val	CCA Pro	GAA Glu 2345	Ile	GTG Val	CCA Pro	TTT Phe	CGC Arg 2350	Leu	ACT Thr	7056
CAT His	AAT Asn	ATG Met 2355	Val	AAT Asn	GGA Gly	ATG Met	GGT Gly 2360	Pro	ATG Met	GGA Gly	ACA Thr	GAG Glu 2365	Gly	CTT Leu	TTT Phe	7104
CGA Arg	AGA Arg 2370	GCA Ala	TGT Cys	GAA Glu	Val	ACA Thr 2375	ATG Met	AGG Arg	CTG Leu	Met .	CGT Arg 2380	Asp	CAG Gln	CGA Arg	GAG Glu	7152
CCT Pro 2385	Leu	ATG Met	AGT Ser	Val	TTA Leu 2390	Lys	ACT Thr	TTT Phe	CTA Leu	CAT His 2395	Asp	CCT Pro	CTT Leu	GTG Val	GAA Glu 2400	7200

-174-

					Lys	GGG Gly				Ala					Thr	7248
				Asn		AAG Lys			Thr					Ile		7296
			Gln			ATC Ile		Thr					Thr			7344
		Ser				CAT His 2455	Val					Gln				7392
	Glu					CAG Gln					Trp					7440
TGA	atg <i>i</i>	I AAJ	TATO	AAAT	A GA	TATA	GTTA	ATA	ATCI	AAA	AGTA	AAAA	LAA A	AAAA	AAAAA	7500
AA																7502

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2480 amino acids (B) TYPE: amino acid

 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Met Gly His Ala Val Glu Trp Pro Val Val Met Ser Arg Phe Leu Ser

Gln Leu Asp Glu His Met Gly Tyr Leu Gln Ser Ala Pro Leu Gln Leu

Met Ser Met Gln Asn Leu Glu Phe Ile Glu Val Thr Leu Leu Met Val

Leu Thr Arg Ile Ile Ala Ile Val Phe Phe Arg Arg Gln Glu Leu Leu

Leu Trp Gln Ile Gly Cys Val Leu Leu Glu Tyr Gly Ser Pro Lys Ile 65 70 75 80

Lys Ser Leu Ala Ile Ser Phe Leu Thr Glu Leu Phe Gln Leu Gly Gly

Leu Pro Ala Gln Pro Ala Ser Thr Phe Phe Ser Ser Phe Leu Glu Leu 100 105

Leu Lys His Leu Val Glu Met Asp Thr Asp Gln Leu Lys Leu Tyr Glu

Glu Pro Leu Ser Lys Leu Ile Lys Thr Leu Phe Pro Phe Glu Ala Glu 135

Ala Tyr Arg Asn Ile Glu Pro Val Tyr Leu Asn Met Leu Leu Glu Lys 155

Leu Cys Val Met Phe Glu Asp Gly Val Leu Met Arg Leu Lys Ser Asp Leu Leu Lys Ala Ala Leu Cys His Leu Leu Gln Tyr Phe Leu Lys Phe Val Pro Ala Gly Tyr Glu Ser Ala Leu Gln Val Arg Lys Val Tyr Val 200 Arg Asn Ile Cys Lys Ala Leu Leu Asp Val Leu Gly Ile Glu Val Asp Ala Glu Tyr Leu Leu Gly Pro Leu Tyr Ala Ala Leu Lys Met Glu Ser Met Glu Ile Ile Glu Glu Ile Gln Cys Gln Thr Gln Glu Asn Leu Ser Ser Asn Ser Asp Gly Ile Ser Pro Lys Arg Arg Leu Ser Ser Ser Leu Asn Pro Ser Lys Arg Ala Pro Lys Gln Thr Glu Glu Ile Lys His Val Asp Met Asn Gln Lys Ser Ile Leu Trp Ser Ala Leu Lys Gln 295 Lys Ala Glu Ser Leu Gln Ile Ser Leu Glu Tyr Ser Gly Leu Lys Asn 305 310 Pro Val Ile Glu Met Leu Glu Gly Ile Ala Val Val Leu Gln Leu Thr Ala Leu Cys Thr Val His Cys Ser His Gln Asn Met Asn Cys Arg Thr 345 Phe Lys Asp Cys Gln His Lys Ser Lys Lys Pro Ser Val Val Ile 360 Thr Trp Met Ser Leu Asp Phe Tyr Thr Lys Val Leu Lys Ser Cys Arg Ser Leu Leu Glu Ser Val Gln Lys Leu Asp Leu Glu Ala Thr Ile Asp 395 Lys Val Val Lys Ile Tyr Asp Ala Leu Ile Tyr Met Gln Val Asn Ser Ser Phe Glu Asp His Ile Leu Glu Asp Leu Cys Gly Met Leu Ser Leu Pro Trp Ile Tyr Ser His Ser Asp Asp Gly Cys Leu Lys Leu Thr Thr Phe Ala Ala Asn Leu Leu Thr Leu Ser Cys Arg Ile Ser Asp Ser Tyr 455 Ser Pro Gln Ala Gln Ser Arg Cys Val Phe Leu Leu Thr Leu Phe Pro 475 Arg Arg Ile Phe Leu Glu Trp Arg Thr Ala Val Tyr Asn Trp Ala Leu Gln Ser Ser His Glu Val Ile Arg Ala Ser Cys Val Ser Gly Phe Phe 500 505

Ile	Leu	Leu 515	Gln	Gln	Gln	Asn	Ser 520	Cys	Asn	Arg	Val	Pro 525	Lys	Ile	Leu
Ile	Asp 530	Lys	Val	Lys	Asp	Asp 535	Ser	Asp	Ile	Val	Lys 540	Lys	Glu	Phe	Ala
Ser 545	Ile	Leu	Gly	Gln	Leu 550	Val	Cys	Thr	Leu	His 555	Gly	Met	Phe	Tyr	Leu 560
Thr	Ser	Ser	Leu	Thr 565	Glu	Pro	Phe	Ser	Glu 570	His	Gly	His	Val	Asp 575	Leu
Phe	Cys	Arg	Asn 580	Leu	Lys	Ala	Thr	Ser 585	Gln	His	Glu	Cys	Ser 590	Ser	Ser
Gln	Leu	Lys 595	Ala	Ser	Val	Cys	Lys 600	Pro	Phe	Leu	Phe	Leu 605	Leu	Lys	Lys
Lys	Ile 610	Pro	Ser	Pro	Val.	Lys 615	Leu	Ala	Phe	Ile	Asp 620	Asn	Leu	His	His
Leu 625	Cys	Lys	His	Leu	Asp 630	Phe	Arg	Glu	Asp	Glu 635	Thr	Asp	Val	Lys	Ala 640
Val	Leu	Gly	Thr	Leu 645	Leu	Asn	Leu	Met	Glu 650	Asp	Pro	Asp	Lys	Asp 655	Val
Arg	Val	Ala	Phe 660	Ser	Gly	Asn	Ile	Lys 665	His	Ile	Leu	Glu	Ser 670	Leu	Asp
Ser	Glu	Asp 675	Gly	Phe	Ile	Lys	Glu 680	Leu	Phe	Val	Leu	Arg 685	Met	Lys	Glu
Ala	Tyr 690	Thr	His	Ala	Gln	Ile 695	Ser	Arg	Asn	Asn	Glu 700	Leu	Lys	Asp	Thr
Leu 705	Ile	Leu	Thr	Thr	Gly 710	Asp	Ile	Gly	Arg	Ala 715	Ala	Lys	Gly	Asp	Leu 720
Val	Pro	Phe	Ala	Leu 725	Leu	His	Leu	Leu	His 730	Cys	Leu	Leu	Ser	Lys 735	Ser
Ala	Ser	Val	Ser 740	Gly	Ala	Ala	Tyr	Thr 745	Glu	Ile	Arg	Ala	Leu 750	Val	Ala
Ala	Lys	Ser 755	Val	Lys	Leu	Gln	Ser 760	Phe	Phe	Ser	Gln	Tyr 765	Lys	Lys	Pro
Ile	Cys 770	Gln	Phe	Leu	Val	Glu 775	Ser	Leu	His	Ser	Ser 780	Gln	Met	Thr	Ala
Leu 785	Pro	Asn	Thr	Pro	Cys 790	Gln	Asn	Ala	Asp	Val 795	Arg	Lys	Gln	Asp	Val 800
Ala	His	Gln	Arg	Glu 805	Met	Ala	Leu	Asn	Thr 810	Leu	Ser	Glu	Ile	Ala 815	Asn
Val	Phe	qaA	Phe 820	Pro	Asp	Leu	Asn	Arg 825	Phe	Leu	Thr	Arg	Thr 830	Leu	Gln
Val	Leu	Leu 835	Pro	Asp	Leu	Ala	Ala 840	Lys	Ala	Ser	Pro	Ala 845	Ala	Ser	Ala
Leu	Ile 850	Arg	Thr	Leu	Gly	Lys 855	Gln	Leu	Asn	Val.	Asn 860	Arg	Arg	Glu	Ile

Leu	Ile	Asn	Asn	Phe	Lys	Tyr	· Ile	Phe	Ser	His	Leu	Val	Cvs	Ser	Cys
865	i				870)				875	i		-		880
Ser	Lys	: Asp	Glu	Leu 885		Arg	Ala	Leu	His 890		Leu	Lys	Asn	895	Thr
Glu	Ile	Glu	900		Ser	Leu	Leu	Arg 905		Asp	Phe	Glm	Gly 910		His
Asn	Glu	Leu 915	Leu	Leu	Arg	Ile	Gly 920		His	Tyr	Gln	Gln 925		Phe	Asn
Gly	Leu 930	Ser	Ile	Leu	Ala	Ser 935	Phe	Ala	Ser	Ser	Asp 940	Asp	Pro	Tyr	Gln
Gly 945	Pro	Arg	Asp	Ile	11e 950	Ser	Pro	Glu	Leu	Met 955	Ala	Asp	Tyr	Leu	Gln 960
Pro	Lys	Leu	Leu	Gly 965	Ile	Leu	Ala	Phe	Phe 970	Asn	Met	Gln	Leu	Leu 975	Ser
Ser	Ser	Val	Gly 980	Ile	Glu	Asp	Lys	Lys 985	Met	Ala	Leu	Asn	Ser 990	Leu	Met
Ser	Leu	Met 995	Lys	Leu	Met	Gly	Pro 100		His	Val	Ser	Ser 100		Arg	Val
Lys	Met 101	Met 0	Thr	Thr	Leu	Arg 101	Thr 5	Gly	Leu	Arg	Phe 102		Asp	qaA	Phe
Pro 102	Glu S	Leu	Cys	Cys	Arg 103		Trp	Asp	Cys	Phe 103		Arg	Cys	Leu	Asp 1040
His	Ala	Cys	Leu	Gly 1049		Leu	Leu	Ser	His 105		Ile	Val	Ala	Leu 105	
Pro	Leu	Ile	His 106	Ile O	Gln	Pro	Lys	Glu 1065		Ala	Ala	Ile	Phe	His O	Tyr
Leu	Ile	Ile 1075	Glu 5	Asn	Arg	Asp	Ala 1080		Gln	Asp	Phe	Leu 108		Glu	Ile
Tyr	Phe 109(Leu)	Pro	Asp	His	Pro 1095	Glu 5	Leu	Lys	Lys	Ile 1100		Ala	Val	Leu
Gln 1105	Glu 5	Tyr	Arg	Lys	Glu 1110	Thr	Ser	Glu	Ser	Thr		Leu	Gln	Thr	Thr 1120
Leu	Gln	Leu	Ser	Met 1125	Lys	Ala	Ile	Gln	His 1130	Glu)	Asn	Val	Asp	Val 1135	
Ile	His	Ala	Leu 1140	Thr	Ser	Leu	Lys	Glu 1145	Thr	Leu	Tyr	Lys	Asn 1150	Gln	Glu
Lys	Leu	Ile 1155	Lys	Tyr	Ala	Thr	Asp 1160		Glu	Thr	Val	Glu 1165		Ile	Ile
Ser	Gln 1170	Leu)	Val	Thr	Val	Leu 1175	Leu	Lys	Gly	Cys	Gln 1180		Ala	Asn	Ser
31n 1185	Ala	Arg	Leu	Leu	Cys 1190	Gly	Glu	Cys	Leu	Gly 1195		Leu	Gly	Ala	Ile 1200
Asp	Pro	Gly	Arg	Leu 1205	Asp	Phe	Ser	Thr	Thr 1210		Thr	Gln	Gly	Lys 1215	

. 11

- Phe Thr Phe Val Thr Gly Val Glu Asp Ser Ser Phe Ala Tyr Gly Leu 1220 1225 1230
- Leu Met Glu Leu Thr Arg Ala Tyr Leu Ala Tyr Ala Asp Asn Ser Arg 1235 1240 1245
- Ala Gln Asp Ser Ala Ala Tyr Ala Ile Gln Glu Leu Leu Ser Ile Tyr 1250 1260
- Asp Cys Arg Glu Met Glu Thr Asn Gly Pro Gly His Gln Leu Trp Arg 1265 1270 1275 1280
- Arg Phe Pro Glu His Val Arg Glu Ile Leu Glu Pro His Leu Asn Thr 1285 1290 1295
- Arg Tyr Lys Ser Ser Gln Lys Ser Thr Asp Trp Ser Gly Val Lys Lys 1300 1305 1310
- Pro Ile Tyr Leu Ser Lys Leu Gly Ser Asn Phe Ala Glu Trp Ser Ala 1315 1320 1325
- Ser Trp Ala Gly Tyr Leu Ile Thr Lys Val Arg His Asp Leu Ala Ser 1330 1340
- Lys Ile Phe Thr Cys Cys Ser Ile Met Met Lys His Asp Phe Lys Val 1345 1350 1355 1360
- Thr Ile Tyr Leu Leu Pro His Ile Leu Val Tyr Val Leu Leu Gly Cys 1365 1370 1375
- Asn Gln Glu Asp Gln Gln Glu Val Tyr Ala Glu Ile Met Ala Val Leu 1380 1385 1390
- Lys His Asp Asp Gln His Thr Ile Asn Thr Gln Asp Ile Ala Ser Asp 1395 1400 1405
- Leu Cys Gln Leu Ser Thr Gln Thr Val Phe Ser Met Leu Asp His Leu 1410 1415 1420
- Thr Gln Trp Ala Arg His Lys Phe Gln Ala Leu Lys Ala Glu Lys Cys 1425 1430 1435 1440
- Pro His Ser Lys Ser Asn Arg Asn Lys Val Asp Ser Met Val Ser Thr 1445 1450 1455
- Val Asp Tyr Glu Asp Tyr Gln Ser Val Thr Arg Phe Leu Asp Leu Ile 1460 1465 1470
- Pro Gln Asp Thr Leu Ala Val Ala Ser Phe Arg Ser Lys Ala Tyr Thr 1475 1480 1485
- Arg Ala Val Met His Phe Glu Ser Phe Ile Thr Glu Lys Lys Gln Asn 1490 1495 1500
- Ile Gln Glu His Leu Gly Phe Leu Gln Lys Leu Tyr Ala Ala Met His 1505 1510 1515 1520
- Glu Pro Asp Gly Val Ala Gly Val Ser Ala Ile Arg Lys Ala Glu Pro 1525 1530 1535
- Ser Leu Lys Glu Gln Ile Leu Glu His Glu Ser Leu Gly Leu Leu Arg 1540 1545 1550
- Asp Ala Thr Ala Cys Tyr Asp Arg Ala Ile Gln Leu Glu Pro Asp Gln 1555 1560 1565

- Ile Ile His Tyr His Gly Val Val Lys Ser Met Leu Gly Leu Gly Gln 1570 1580
- Leu Ser Thr Val Ile Thr Gln Val Asn Gly Val His Ala Asn Arg Ser 1585 1590 1595 1600
- Glu Trp Thr Asp Glu Leu Asn Thr Tyr Arg Val Glu Ala Ala Trp Lys 1605 1610 1615
- Leu Ser Gln Trp Asp Leu Val Glu Asn Tyr Leu Ala Ala Asp Gly Lys 1620 1625 1630
- Ser Thr Thr Trp Ser Val Arg Leu Gly Gln Leu Leu Ser Ala Lys 1635 1640 1645
- Lys Arg Asp Ile Thr Ala Phe Tyr Asp Ser Leu Lys Leu Val Arg Ala 1650 1655 1660
- Glu Gln Ile Val Pro Leu Ser Ala Ala Ser Phe Glu Arg Gly Ser Tyr 1665 1670 1675 1680
- Gln Arg Gly Tyr Glu Tyr Ile Val Arg Leu His Met Leu Cys Glu Leu 1685 1690 1695
- Glu His Ser Ile Lys Pro Leu Phe Gln His Ser Pro Gly Asp Ser Ser 1700 1705 1710
- Gln Glu Asp Ser Leu Asn Trp Val Ala Arg Leu Glu Met Thr Gln Asn 1715 1720 1725
- Ser Tyr Arg Ala Lys Glu Pro Ile Leu Ala Leu Arg Arg Ala Leu Leu 1730 1740
- Ser Leu Asn Lys Arg Pro Asp Tyr Asn Glu Met Val Gly Glu Cys Trp 1745 1750 1755 1760
- Leu Gln Ser Ala Arg Val Ala Arg Lys Ala Gly His His Gln Thr Ala 1765 1770 1775
- Tyr Asn Ala Leu Leu Asn Ala Gly Glu Ser Arg Leu Ala Glu Leu Tyr 1780 1785 1790
- Val Glu Arg Ala Lys Trp Leu Trp Ser Lys Gly Asp Val His Gln Ala 1795 1800 1805
- Leu Ile Val Leu Gln Lys Gly Val Glu Leu Cys Phe Pro Glu Asn Glu 1810 1815 1820
- Thr Pro Pro Glu Gly Lys Asn Met Leu Ile His Gly Arg Ala Met Leu 1825 1830 1835 1840
- Leu Val Gly Arg Phe Met Glu Glu Thr Ala Asn Phe Glu Ser Asn Ala 1845 1850 1855
- Ile Met Lys Lys Tyr Lys Asp Val Thr Ala Cys Leu Pro Glu Trp Glu 1860 1865 1870
- Asp Gly His Phe Tyr Leu Ala Lys Tyr Tyr Asp Lys Leu Met Pro Met 1875 1880 1885
- Val Thr Asp Asn Lys Met Glu Lys Gln Gly Asp Leu Ile Arg Tyr Ile 1890 1895 1900
- Val Leu His Phe Gly Arg Ser Leu Gln Tyr Gly Asn Gln Phe Ile Tyr 1905 1910 1915 1920

- Gln Ser Met Pro Arg Met Leu Thr Leu Trp Leu Asp Tyr Gly Thr Lys 1925 1930 1935
- Ala Tyr Glu Trp Glu Lys Ala Gly Arg Ser Asp Arg Val Gln Met Arg 1940 1945 1950
- Asn Asp Leu Gly Lys Ile Asn Lys Val Ile Thr Glu His Thr Asn Tyr 1955 1960 1965
- Leu Ala Pro Tyr Gln Phe Leu Thr Ala Phe Ser Gln Leu Ile Ser Arg 1970 1975 1980
- Ile Cys His Ser His Asp Glu Val Phe Val Val Leu Met Glu Ile Ile 1985 1990 1995 2000
- Ala Lys Val Phe Leu Ala Tyr Pro Gln Gln Ala Met Trp Met Met Thr 2005 2010 2015
- Ala Val Ser Lys Ser Ser Tyr Pro Met Arg Val Asn Arg Cys Lys Glu 2020 2025 2030
- Ile Leu Asn Lys Ala Ile His Met Lys Lys Ser Leu Glu Lys Phe Val 2035 2040 2045
- Gly Asp Ala Thr Arg Leu Thr Asp Lys Leu Leu Glu Leu Cys Asn Lys 2050 2055 2060
- Pro Val Asp Gly Ser Ser Ser Thr Leu Ser Met Ser Thr His Phe Lys 2065 2070 2075 2080
- Met Leu Lys Lys Leu Val Glu Glu Ala Thr Phe Ser Glu Ile Leu Ile 2085 2090 2095
- Pro Leu Gln Ser Val Met Ile Pro Thr Leu Pro Ser Ile Leu Gly Thr 2100 2105 2110
- His Ala Asn His Ala Ser His Glu Pro Phe Pro Gly His Trp Ala Tyr 2115 2120 2125
- Ile Ala Gly Phe Asp Asp Met Val Glu Ile Leu Ala Ser Leu Gln Lys 2130 2135 2140
- Pro Lys Lys Ile Ser Leu Lys Gly Ser Asp Gly Lys Phe Tyr Ile Met 2145 2150 2155 2160
- Met Cys Lys Pro Lys Asp Asp Leu Arg Lys Asp Cys Arg Leu Met Glu 2165 2170 2175
- Phe Asn Ser Leu Ile Asn Lys Cys Leu Arg Lys Asp Ala Glu Ser Arg 2180 2185 2190
- Arg Arg Glu Leu His Ile Arg Thr Tyr Ala Val Ile Pro Leu Asn Asp 2195 2200 2205
- Glu Cys Gly Ile Ile Glu Trp Val Asn Asn Thr Ala Gly Leu Arg Pro 2210 2215 2220
- Ile Leu Thr Lys Leu Tyr Lys Glu Lys Gly Val Tyr Met Thr Gly Lys 2225 2230 2235 2240
- Glu Leu Arg Gln Cys Met Leu Pro Lys Ser Ala Ala Leu Ser Glu Lys 2245 2250 2255
- Leu Lys Val Phe Arg Glu Phe Leu Leu Pro Arg His Pro Pro Ile Phe 2260 2265 2270

60

120

180

-181-

His Glu Trp Phe Leu Arg Thr Phe Pro Asp Pro Thr Ser Trp Tyr Ser

		2275	5				2280)				2285	5		
Ser	Arg 2290		Ala	Tyr	Cys	Arg 2295		Thr	Ala	Val	Met 2300		Met	Val	Gly
Tyr 2305	Ile	Leu	Gly	Leu	Gly 2310		Arg	His	Gly	Glu 2315		Ile	Leu	Phe	Asp 2320
Ser	Leu	Thr	Gly	Glu 2325		Val	His	Val	Asp 2330		Asn	аұЭ	Leu	Phe 2335	
Lys	Gly	Glu	Thr 2340		Glu	Val	Pro	Glu 2345		Val	Pro	Phe	Arg 2350		Thr
His	Asn	Met 2355		Asn	Gly	Met	Gly 2360		Met	Gly	Thr	Glu 2365		Leu	Phe
Arg	Arg 2370		Cys	Glu	Val	Thr 2375		Arg	Leu	Met	Arg 2380		Gln	Arg	Glu
Pro 2385	Leu	Met	Ser	Val	Leu 2390	-	Thr	Phe	Leu	His 2395	-	Pro	Leu	Val	Glu 2400
Trp	Ser	Lys	Pro	Val 2405		Gly	His	Ser	Lys 2410		Pro	Leu	Asn	Glu 2415	
Gly	Glu	Val	Val 2420		Glu	Lys	Ala	Lys 2425		His	Val	Leu	Asp 2430		Glu
Gln	Arg	Leu 2435		Gly	Val	Ile	Lys 2440		Arg	Asn	Arg	Val 2445		Gly	Leu
Pro	Leu 2450		Ile	Glu	Gly	His 2455		His	Tyr	Leu	Ile 2460		Glu	Ala	Thr
Asp 2465	Glu S	Asn	Leu	Leu	Cys 2470		Met	Tyr	Leu	Gly 2475	_	Thr	Pro	Tyr	Met 2480
(2)	INFO	RMAT	NOI	FOR	SEQ	ID N	10 : 34	:							
	(i)	(<i>E</i> (C	A) LE B) TY C) ST	E CH NGTH PE: RAND	: 93 nucl	85 b eic SS:	ase ació sino	pair I	rs						
	(ii)	MOI	ECUI	E TY	PE:	CDNA									
	(v)	FRA	GMEN	T TY	PE:	line	ar								
	(ix)	(P		: ME/K CATI			.935	57							
	(xi)	SEÇ	UENC	E DE	SCRI	PTIC	N: S	EQ I	מ סו	34:					

GCGAGAGGAG TCGGGATCTG CGCTGCAGCC ACCGCCGCGG TTGATACTAC TTTGACCTTC

CGAGTGCAGT GAGGCATACA TCACAATTTG GAATTATGCA TTGGTTTATC AATTTACTTG

TTTATTGTCA CCCTGCTGCC CAGATATGAC TTCATGAGGA CAGTGATGTG TGTTCTGAAA

TTG	TGAA	CC A	TG A let S	GT C er L	TA G	TA C	TT A eu A 5	AT G .sn A	AT C	TG C eu L	TT A eu I	TC T le C 10	GC T ys C	GC C	GT rg	228
CAA Gln	CTA Leu 15	GAA Glu	CAT His	GAT Asp	AGA Arg	GCT Ala 20	ACA Thr	GAA Glu	CGA Arg	AAG Lys	AAA Lys 25	Glu	GTT Val	GAG Glu	AAA Lys	276
TTT Phe 30	AAG Lys	CGC Arg	CTG Leu	ATT	CGA Arg 35	GAT Asp	CCT Pro	GAA Glu	ACA Thr	ATT Ile 40	AAA Lys	CAT His	CTA Leu	GAT Asp	CGG Arg 45	324
CAT His	TCA Ser	GAT Asp	TCC Ser	AAA Lys 50	Gln	GGA Gly	AAA Lys	TAT Tyr	TTG Leu 55	TAA Asn	TGG Trp	GAT Asp	GCT Ala	GTT Val 60	TTT Phe	372
AGA Arg	TTT Phe	TTA Leu	CAG Gln 65	AAA Lys	TAT Tyr	ATT	CAG Gln	AAA Lys 70	GAA Glu	ACA Thr	GAA Glu	TGT Cys	CTG Leu 75	AGA Arg	ATA Ile	420
GCA Ala	AAA Lys	CCA Pro 80	AAT Asn	GTA Val	TCA Ser	GCC Ala	TCA Ser 85	ACA Thr	CAA Gln	GCC Ala	TCC Ser	AGG Arg 90	CAG Gln	AAA Lys	AAG Lys	468
ATG Met	CAG Gln 95	GAA Glu	ATC Ile	AGT Ser	AGT Ser	TTG Leu 100	GTC Val	AAA Lys	TAC Tyr	TTC Phe	ATC Ile 105	AAA Lys	TGT Cys	GCA Ala	AAC Asn	516
AGA Arg 110	AGA Arg	GCA Ala	CCT Pro	AGG Arg	CTA Leu 115	AAA Lys	TGT Cys	CAA Gln	GAA Glu	CTC Leu 120	TTA Leu	AAT Asn	TAT Tyr	ATC Ile	ATG Met 125	564
GAT Asp	ACA Thr	GTG Val	AAA Lys	GAT Asp 130	TCA Ser	TCT Ser	AAT Asn	GGT Gly	GCT Ala 135	ATT Ile	TAC Tyr	GGA Gly	GCT Ala	GAT Asp 140	TGT Cys	612
AGC Ser	AAC Asn	ATA Ile	CTA Leu 145	CTC Leu	Lys	GAC Asp	ATT Ile	CTT Leu 150	TCT Ser	GTG Val	AGA Arg	AAA Lys	TAC Tyr 155	TGG Trp	TGT Cys	660
GAA Glu	ATA Ile	TCT Ser 160	CAG Gln	CAA Gln	CAG Gln	TGG Trp	TTA Leu 165	GAA Glu	TTG Leu	TTC Phe	TCT Ser	GTG Val 170	TAC Tyr	TTC Phe	AGG Arg	708
CTC Leu	TAT Tyr 175	CTG Leu	AAA Lys	CCT Pro	TCA Ser	CAA Gln 180	GAT Asp	GTT Val	CAT His	AGA Arg	GTT Val 185	TTA Leu	GTG Val	GCT Ala	AGA Arg	756
ATA Ile 190	ATT Ile	CAT His	GCT Ala	GTT Val	ACC Thr 195	AAA Lys	GGA Gly	TGC Cys	TGT Cys	TCT Ser 200	CAG Gln	ACT Thr	GAC Asp	GGA Gly	TTA Leu 205	804
AAT Asn	TCC Ser	AAA Lys	TTT Phe	TTG Leu 210	GAC Asp	TTT Phe	TTT Phe	TCC Ser	AAG Lys 215	GCT Ala	ATT Ile	CAG Gln	TGT Cys	GCG Ala 220	AGA Arg	852
CAA Gln	GAA Glu	AAG Lys	AGC Ser 225	TCT Ser	TCA Ser	GGT Gly	CTA Leu	AAT Asn 230	CAT His	ATC Ile	TTA Leu	GCA Ala	GCT Ala 235	CTT Leu	ACT Thr	900
ATC Ile	TTC Phe	CTC Leu 240	AAG Lys	ACT Thr	TTG Leu	GCT Ala	GTC Val 245	AAC Asn	TTT Phe	CGA Arg	ATT Ile	CGA Arg 250	GTG Val	TGT Cys	GAA Glu	948
TTA Leu	GGA Gly 255	GAT Asp	GAA Glu	ATT Ile	CTT Leu	CCC Pro 260	ACT Thr	TTG Leu	CTT Leu	TAT Tyr	ATT Ile 265	TGG Trp	ACT Thr	CAA Gln	CAT His	996

AGG Arg 270	Leu	AAT Asn	GAT Asp	TCT Ser	TTA Leu 275	Lys	GAA Glu	GTC Val	ATT Ile	ATT Ile 280	Glu	TTA Leu	TTT Phe	CAA Gln	CTG Leu 285	1044
CAA Gln	ATT	TAT Tyr	ATC Ile	CAT His 290	His	CCG Pro	AAA Lys	GGA Gly	GCC Ala 295	AAA Lys	ACC Thr	CAA Gln	GAA Glu	AAA Lys 300	GGT Gly	1092
GCT Ala	TAT	GAA Glu	TCA Ser 305	Thr	AAA Lys	TGG Trp	AGA Arg	AGT Ser 310	Ile	TTA Leu	TAC	AAC Asn	TTA Leu 315	Tyr	GAT Asp	1140
CTG Leu	CTA Leu	GTG Val 320	Asn	GAG Glu	ATA Ile	AGT Ser	CAT His 325	ATA Ile	GGA Gly	AGT Ser	AGA Arg	GGA Gly 330	Lys	TAT Tyr	TCT Ser	1188
TCA Ser	GGA Gly 335	TTT Phe	CGT Arg	AAT Asn	ATT Ile	GCC Ala 340	GTC Val	AAA Lys	GAA Glu	AAT Asn	TTG Leu 345	ATT	GAA Glu	TTG Leu	ATG Met	1236
GCA Ala 350	GAT Asp	ATC Ile	TGT Cys	CAC His	CAG Gln 355	GTT Val	TTT Phe	AAT Asn	GAA Glu	GAT Asp 360	ACC Thr	AGA Arg	TCC Ser	TTG Leu	GAG Glu 365	1284
ATT Ile	TCT Ser	CAA Gln	TCT Ser	TAC Tyr 370	ACT Thr	ACT Thr	ACA Thr	CAA Gln	AGA Arg 375	GAA Glu	TCT Ser	AGT Ser	GAT Asp	TAC Tyr 380	AGT Ser	1332
GTC Val	CCT Pro	TGC Cys	AAA Lys 385	AGG Arg	AAG Lys	AAA Lys	ATA Ile	GAA Glu 390	CTA Leu	GGC Gly	TGG Trp	GAA Glu	GTA Val 395	ATA Ile	AAA Lys	1380
GAT Asp	CAC His	CTT Leu 400	CAG Gln	AAG Lys	TCA Ser	CAG Gln	AAT Asn 405	GAT Asp	TTT Phe	GAT Asp	CTT Leu	GTG Val 410	CCT Pro	TGG Trp	CTA Leu	1428
CAG Gln	ATT Ile 415	GCA Ala	ACC Thr	CAA Gln	TTA Leu	ATA Ile 420	TCA Ser	AAG Lys	TAT Tyr	CCT Pro	GCA Ala 425	AGT Ser	TTA Leu	CCT Pro	AAC Asn	1476
TGT Cys 430	GAG Glu	CTG Leu	TCT Ser	CCA Pro	TTA Leu 435	CTG Leu	ATG Met	ATA Ile	CTA Leu	TCT Ser 440	CAG Gln	CTT Leu	CTA Leu	CCC Pro	CAA Gln 445	1524
CAG Gln	CGA Arg	CAT His	GGG Gly	GAA Glu 450	CGT Arg	ACA Thr	CCA Pro	TAT Tyr	GTG Val 455	TTA Leu	CGA Arg	TGC Cys	CTT Leu	ACG Thr 460	GAA Glu	1572
GTT Val	GCA Ala	TTG Leu	TGT Cys 465	CAA Gln	GAC Asp	AAG Lys	AGG Arg	TCA Ser 470	AAC Asn	CTA Leu	GAA Glu	AGC Ser	TCA Ser 475	CAA Gln	AAG Lys	1620
TCA Ser	GAT Asp	TTA Leu 480	TTA Leu	AAA Lys	CTC Leu	TGG Trp	AAT Asn 485	AAA Lys	ATT Ile	TGG Trp	TGT Cys	ATT Ile 490	ACC Thr	TTT Phe	CGT Arg	1668
GGT Gly	ATA Ile 495	AGT Ser	TCT Ser	GAG Glu	CAA Gln	ATA Ile 500	CAA Gln	GCT Ala	GAA Glu	AAC Asn	TTT Phe 505	GGC Gly	TTA Leu	CTT Leu	GGA Gly	1716
GCC Ala 510	ATA Ile	ATT Ile	CAG Gln	GGT Gly	AGT Ser 515	TTA Leu	GTT Val	GAG Glu	GTT Val	GAC Asp 520	AGA Arg	GAA Glu	TTC Phe	TGG Trp	AAG Lys 525	1764
TTA Leu	TTT Phe	ACT Thr	GGG Gly	TCA Ser 530	GCC Ala	TGC Cys	AGA Arg	CCT Pro	TCA Ser 535	TGT Cys	CCT Pro	GCA Ala	GTA Val	TGC Cys 540	TGT Cys	1812

		TTG Leu														1860
		GAG Glu 560														1908
GAA Glu	TCA Ser 575	ATA Ile	ATG Met	AAA Lys	TGG Trp	CTC Leu 580	TTA Leu	TTC Phe	TAT Tyr	CAG Gln	TTA Leu 585	GAG Glu	GGT Gly	GAC Asp	TTA Leu	1956
		AGC Ser														2004
CTT Leu	GTA Val	CTG Leu	GAG Glu	AAA Lys 610	ATT Ile	CTT Leu	GTG Val	AGT Ser	CTC Leu 615	ACT Thr	ATG Met	AAA Lys	AAC Asn	TGT Cys 620	AAA Lys	2052
		ATG Met														2100
		AAA Lys 640														2148
		ACT Thr														2196
		GAA Glu														2244
		GAA Glu														2292
		AAT Asn														2340
		CGT Arg 720														2388
		GCT Ala														2436
		CTA Leu														2484
		AAT Asn														2532
		ACA Thr														2580
		TCT Ser 800														2628

		Ala													CCA Pro	2676
TTT Phe 830	GAC Asp	CGT Arg	GGA Gly	GAA Glu	GTA Val 835	GAA Glu	TCA Ser	ATG Met	GAA Glu	GAT Asp 840	GAT Asp	ACT Thr	AAT Asn	GGA Gly	AAT Asn 845	2724
					Asp					AAT Asn						2772
CCT Pro	GAT Asp	AGT Ser	AGT Ser 865	GTT Val	AGT Ser	GAT Asp	GCA Ala	AAC Asn 870	GAA Glu	CCT Pro	GGA Gly	GAG Glu	AGC Ser 875	CAA Gln	AGT Ser	2820
ACC Thr	ATA Ile	GGT Gly 880	GCC Ala	ATT	AAT Asn	CCT Pro	TTA Leu 885	GCT Ala	GAA Glu	GAA Glu	TAT Tyr	CTG Leu 890	TCA Ser	AAG Lys	CAA Gln	2868
GAT Asp	CTA Leu 895	CTT Leu	TTC Phe	TTA Leu	GAC Asp	ATG Met 900	CTC Leu	AAG Lys	TTC Phe	TTG Leu	TGT Cys 905	TTG Leu	тст Сув	GTA Val	ACT Thr	2916
ACT Thr 910	GCT Ala	CAG Gln	ACC Thr	AAT Asn	ACT Thr 915	GTG Val	TCC Ser	TTT Phe	AGG Arg	GCA Ala 920	GCT Ala	GAT Asp	ATT Ile	CGG Arg	AGG Arg 925	2964
AAA Lys	TTG Leu	TTA Leu	ATG Met	TTA Leu 930	ATT Ile	GAT Asp	TCT Ser	AGC Ser	ACG Thr 935	CTA Leu	GAA Glu	CCT Pro	ACC Thr	AAA Lys 940	TCC Ser	3012
CTC Leu	CAC His	CTG Leu	CAT His 945	ATG Met	TAT Tyr	CTA Leu	ATG Met	CTT Leu 950	TTA Leu	AAG Lys	GAG Glu	CTT Leu	CCT Pro 955	GGA Gly	GAA Glu	3060
GAG Glu	TAC Tyr	CCC Pro 960	TTG Leu	CCA Pro	ATG Met	GAA Glu	GAT Asp 965	GTT Val	CTT Leu	GAA Glu	CTT Leu	CTG Leu 970	AAA Lys	CCA Pro	CTA Leu	3108
TCC Ser	AAT Asn 975	GTG Val	TGT Cys	TCT Ser	TTG Leu	TAT Tyr 980	CGT Arg	CGT Arg	GAC Asp	CAA Gln	GAT Asp 985	GTT Val	TGT Cys	AAA Lys	ACT Thr	3156
ATT Ile 990	TTA Leu	AAC Asn	CAT His	GTC Val	CTT Leu 995	CAT His	GTA Val	GTG Val	AAA Lys	AAC Asn 1000	Leu	GGT Gly	CAA Gln	AGC Ser	AAT Asn 1005	3204
ATG Met	GAC Asp	TCT Ser	GAG Glu	AAC Asn 1010	Thr	AGG Arg	GAT Asp	GCT Ala	CAA Gln 1015	GGA Gly	CAG Gln	TTT Phe	CTT Leu	ACA Thr 1020	Val	3252
ATT Ile	GGA Gly	GCA Ala	TTT Phe 1025	Trp	CAT His	CTA Leu	ACA Thr	AAG Lys 1030	Glu	AGG Arg	AAA Lys	TAT Tyr	ATA Ile 1035	Phe	TCT Ser	3300
GTA Val	AGA Arg	ATG Met 1040	Ala	CTA Leu	GTA Val	AAT Asn	TGC Cys 1045	Leu	AAA Lys	ACT Thr	TTG Leu	CTT Leu 1050	Glu	GCT Ala	GAT Asp	3348
CCT Pro	TAT Tyr 1055	Ser	AAA Lys	TGG Trp	GCC Ala	ATT Ile 1060	Leu	AAT Asn	GTA Val	ATG Met	GGA Gly 1065	Lys	GAC Asp	TTT Phe	CCT Pro	3396
GTA Val 1070	Asn	GAA Glu	GTA Val	TTT Phe	ACA Thr 1075	Gln	TTT Phe	CTT Leu	GCT Ala	GAC Asp 1080	Asn	CAT His	CAC His	CAA Gln	GTT Val 1085	3444

CGC ATG TTG GC Arg Met Leu Al	CT GCA GAG TCA AT la Ala Glu Ser Il 1090	C AAT AGA TTG T e Asn Arg Leu Pl 1095	TC CAG GAC ACG he Gln Asp Thr 110	Lys
Gly Asp Ser Se	CC AGG TTA CTG AA er Arg Leu Leu Ly 105	A GCA CTT CCT T S Ala Leu Pro Le 1110	TG AAG CTT CAG eu Lys Leu Gln 1115	CAA 3540 Gln
	AA AAT GCA TAC TTO lu Asn Ala Tyr Leo li	ı Lys Ala Gln G		
ATG TCC CAT AG Met Ser His Se 1135	ET GCT GAG AAC CC er Ala Glu Asn Pro 1140	o Glu Thr Leu As	AT GAA ATT TAT sp Glu Ile Tyr 145	AAT 3636 Asn
AGA AAA TCT GT Arg Lys Ser Va 1150	TT TTA CTG ACG TTG al Leu Leu Thr Leu 1155	G ATA GCT GTG GT I Ile Ala Val Va 1160	TT TTA TCC TGT al Leu Ser Cys	AGC 3684 Ser 1165
CCT ATC TGC GA Pro Ile Cys Gl	AA AAA CAG GCT TTC Lu Lys Gln Ala Leo 1170	G TTT GCC CTG TO 1 Phe Ala Leu Cy 1175	GT AAA TCT GTG ys Lys Ser Val 118	Lys
Glu Asn Gly Le	TA GAA CCT CAC CT eu Glu Pro His Lei 185	GTG AAA AAG GT Val Lys Lys Va 1190	TT TTA GAG AAA al Leu Glu Lys 1195	GTT 3780 Val
	TT GGA TAT AGA CG ne Gly Tyr Arg Arg 120	Leu Glu Asp Ph		
	rG GTT TTG GAA TGG eu Val Leu Glu Trṛ 1220	Leu Asn Leu Gl		
	CT TTT CCT TTT AT: er Phe Pro Phe Ile 1235			
	GA TCT TGT TAT AAC TG Ser Cys Tyr Lys 1250			Ile
Arg Ser His Ph	TT GAT GAG GTG AAG ne Asp Glu Val Ly: 265			
	GT CTT CTA ACA GAG er Leu Leu Thr Asp 128	Cys Phe Pro Ly		
	AT TTT GCC TAT GAG or Phe Ala Tyr Glu 1300	Gly Thr Arg As		
CAG CAA AGA GA Gln Gln Arg Gl 1310	AG ACT GCT ACC AAC Lu Thr Ala Thr Lys 1315	GTC TAT GAT AT Val Tyr Asp Me 1320	TG CTT AAA AGT et Leu Lys Ser	GAA 4164 Glu 1325
AAC TTA TTG GG Asn Leu Leu Gl	GA AAA CAG ATT GA Ly Lys Gln Ile Asp 1330	C CAC TTA TTC AT His Leu Phe II 1335	TT AGT AAT TTA le Ser Asn Leu 134	Pro
Glu Ile Val Va	TG GAG TTA TTG ATG al Glu Leu Leu Met 845	ACG TTA CAT GAT Thr Leu His GI 1350	AG CCA GCA AAT lu Pro Ala Asn 1355	TCT 4260 Ser

AGT Ser	GCC Ala	AGT Ser 136	Gln	AGC Ser	ACT Thi	GAC Asp	CTC Leu 136	Cys	GAC Asp	TTT Phe	TCA Ser	GGG Gly	Asp	TTG	GAT Asp	4309
CCT Pro	GCT Ala 137	Pro	AAT Asn	CCA Pro	CCI Pro	CAT His 138	Phe	CCA Pro	TCG Ser	CAT	GTG Val	Ile	AAA Lys	GCA Ala	ACA Thr	4356
TTT Phe 139	Ala	TAT Tyr	ATC Ile	AGC Ser	AAT Asn 139	TGT Cys	CAT His	AAA Lys	ACC Thr	AAG Lys 140	Leu	AAA Lys	AGC Ser	ATT Ile	TTA Leu 1405	4404
GAA Glu	ATT	CTT	TCC Ser	AAA Lys 141	Ser	CCT Pro	GAT Asp	TCC Ser	TAT Tyr 141	Gln	AAA Lys	ATT	CTT Leu	CTT Leu 142	Ala	4452
ATA Ile	TGT Cys	GAG Glu	CAA Gln 142	Ala	GCT Ala	GAA Glu	ACA Thr	AAT Asn 143	Asn	GTT Val	TAT Tyr	AAG Lys	AAG Lys 143	His	AGA Arg	4500
ATT Ile	CTT Leu	AAA Lys 144	Ile	TAT Tyr	CAC	CTG Leu	TTT Phe 144	Val	AGT Ser	TTA Leu	TTA Leu	CTG Leu 145	Lys	GAT Asp	ATA Ile	4548
AAA Lys	AGT Ser 145	Gly	TTA Leu	GGA Gly	GGA Gly	GCT Ala 1460	Trp	GCC Ala	TTT Phe	GTT Val	CTT Leu 146	Arg	GAC Asp	GTT Val	ATT Ile	4596
TAT Tyr 147	Thr	TTG Leu	ATT	CAC His	TAT Tyr 147	ATC Ile 5	AAC Asn	CAA Gln	AGG Arg	CCT Pro 148	Ser	TGT Cys	ATC Ile	ATG Met	GAT Asp 1485	4644
GTG Val	TCA Ser	TTA Leu	CGT Arg	AGC Ser 1490	Phe	TCC Ser	CTT Leu	TGT Cys	TGT Cys 1495	Asp	TTA Leu	TTA Leu	AGT Ser	CAG Gln 1500	Val	4692
TGC Cys	CAG Gln	ACA Thr	GCC Ala 1505	Val	ACT Thr	TAC Tyr	TGT Cys	AAG Lys 1510	Asp	GCT Ala	CTA Leu	GAA Glu	AAC Asn 1515	His	CTT Leu	4740
CAT His	GTT Val	ATT Ile 1520	Val	GGT Gly	ACA Thr	CTT Leu	ATA Ile 1525	Pro	CTT Leu	GTG Val	TAT Tyr	GAG Glu 1530	Gln	GTG Val	GAG Glu	4788.
GTT Val	CAG Gln 1535	Lys	CAG Gln	GTA Val	TTG Leu	GAC Asp 1540	Leu	TTG Leu	AAA Lys	TAC Tyr	TTA Leu 1545	Val	ATA Ile	GAT Asp	AAC Asn	4836
AAG Lys 1550	qaA	AAT Asn	GAA Glu	AAC Asn	CTC Leu 1555	TAT Tyr	ATC Ile	ACG Thr	ATT Ile	AAG Lys 1560	Leu	TTA Leu	GAT Asp	CCT Pro	TTT Phe 1565	4884
CCT Pro	GAC Asp	CAT His	GTT Val	GTT Val 1570	Phe	AAG Lys	GAT Asp	TTG Leu	CGT Arg 1575	Ile	ACT Thr	CAG Gln	CAA Gln	AAA Lys 1580	Ile	4932
AAA Lys	TAC Tyr	AGT Ser	AGA Arg 1585	Gly	CCC Pro	TTT Phe	Ser	CTC Leu 1590	Leu	GAG Glu	GAA Glu	ATT Ile	AAC Asn 1595	His	TTT Phe	4980
CTC Leu	TCA Ser	GTA Val 1600	Ser	GTT Val	TAT Tyr	GAT Asp	GCA Ala 1605	Leu	CCA Pro	TTG Leu	ACA Thr	AGA Arg 1610	Leu	GAA Glu	GGA Gly	5028
Leu	AAG Lys 1615	Asp	CTT Leu	CGA . Arg .	AGA Arg	CAA Gln 1620	CTG (Leu (GAA Glu	CTA Leu	CAT His	AAA Lys 1625	Asp	CAG Gln	ATG Met	GTG Val	5076

GAC ATT ATG AGA Asp Ile Met Arg 1630					
AAA CTA GTT GTC Lys Leu Val Val	AAT TTG TTG (Asn Leu Leu (1650	CAG TTA TCC Gln Leu Ser 1655	Lys Met Ala 1	ATA AAC (le Asn 1660	His
ACT GGT GAA AAA Thr Gly Glu Lys 1669	Glu Val Leu C		Gly Ser Cys I		
GTG GGT CCT ATA Val Gly Pro Ile 1680	Asp Phe Ser 7				
GCA TCT TAT ACC Ala Ser Tyr Thr 1695	AAG GCC CTT I Lys Ala Leu I 1700	AAG TTA TTT Lys Leu Phe	GAA GAT AAA G Glu Asp Lys G 1705	AA CTT	CAG 5316 Gln
TGG ACC TTC ATA Trp Thr Phe Ile 1710	ATG CTG ACC T Met Leu Thr 7 1715	Tyr Leu Asn	AAC ACA CTG G Asn Thr Leu V 1720	al Glu	GAT 5364 Asp 1725
TGT GTC AAA GTT Cys Val Lys Val			Cys Leu Lys A		
GCC ACA AAG ACT Ala Thr Lys Thr 1745	Gly His Ser B		Ile Tyr Lys M		
GAT CCA ATG CTG Asp Pro Met Leu 1760	Ala Tyr Leu G	CAG CCT TTT . Gln Pro Phe . 1765	AGA ACA TCA A Arg Thr Ser A 1770	GA AAA rg Lys	AAG 5508 Lys
TTT TTA GAA GTA Phe Leu Glu Val 1775					
GAT GAT ATA AAT Asp Asp Ile Asn 1790		Pro Leu Ser		sp Ile	
ATA AAG ACA CTG Ile Lys Thr Leu			Ser Gly Gly T		Cys
GAA ATT CTT CAA Glu Ile Leu Gln 1829	Leu Leu Lys F		Glu Val Lys T		
TGT CAG ACT GTA Cys Gln Thr Val 1840	Leu Pro Tyr I				
ACA AAT GAA TCA Thr Asn Glu Ser 1855		Leu Leu Ser			
TTC ACC AGC TGT Phe Thr Ser Cys 1870	•	Phe Ser Gln			
CCT GCA AAC TTG Pro Ala Asn Leu			Phe Phe Arg C		Leu

GAT Asp	AAA Lys	AAA Lys	TCA Ser 190	Gln	AGA Arg	ACA Thr	ATG Met	CTI Leu 191	Ala	GTT Val	GTC Val	GAC Asp	TAC Tyr 191	Met	G AGA : Arg	5940	
AGA Arg	CAA Gln	AAG Lys 192	Arg	CCT Pro	TCT Ser	TCA Ser	GGA Gly 192	Thr	ATT	TTT Phe	AAT Asn	GAT Asp 193	Ala	TTC Phe	TGG Trp	5988	
CTG Leu	GAT Asp 193	Leu	TAA naA	TAT Tyr	CTA Leu	GAA Glu 194	Val	GCC Ala	AAG Lys	GTA Val	GCT Ala 194	Gln	TCT Ser	TGT Cys	GCT Ala	6036	
GCT Ala 1950	His	TTT Phe	ACA Thr	GCT Ala	TTA Leu 195	Leu	TAT Tyr	GCA Ala	GAA Glu	ATC Ile 196	Tyr	GCA Ala	GAT Asp	AAG Lys	AAA Lys 1965	6084	
AGT Ser	ATG Met	GAT Asp	GAT Asp	CAA Gln 197	Glu	AAA Lys	AGA Arg	AGT Ser	CTT Leu 197	Ala	TTT Phe	GAA Glu	GAA Glu	GGA Gly 198	AGC Ser 0	6132	
CAG Gln	AGT Ser	ACA Thr	ACT Thr 1989	Ile	TCT Ser	AGC Ser	TTG Leu	AGT Ser 199	Glu	AAA Lys	AGT Ser	AAA Lys	GAA Glu 199	Glu	ACT Thr	6180	
GGA Gly	ATA Ile	AGT Ser 2000	Leu	CAG Gln	GAT Asp	CTT Leu	CTC Leu 200	Leu	GAA Glu	ATC Ile	TAC Tyr	AGA Arg 201	Ser	ATA Ile	GGG Gly	6228	
GAG Glu	CCA Pro 2019	Asp	AGT Ser	TTG Leu	TAT Tyr	GGC Gly 2020	Cys	GGT Gly	GGA Gly	GGG Gly	AAG Lys 202	Met	TTA Leu	CAA Gln	CCC Pro	6276	
ATT Ile ' 2030	Thr	AGA Arg	CTA Leu	CGA Arg	ACA Thr 2035	Tyr	GAA Glu	CAC His	GAA Glu	GCA Ala 2040	Met	TGG Trp	GGC Gly	AAA Lys	GCC Ala 2045	6324	
CTA (GTA Val	ACA Thr	TAT Tyr	GAC Asp 2050	Leu	GAA Glu	ACA Thr	GCA Ala	ATC 11e 2055	Pro	TCA Ser	TCA Ser	ACA Thr	CGC Arg 206	Gln	6372	
GCA (GGA Gly	ATC Ile	ATT Ile 2065	Gln	GCC Ala	TTG Leu	CAG Gln	AAT Asn 2070	Leu	GGA Gly	CTC Leu	TGC Cys	CAT His 2075	Ile	CTT Leu	6420	
TCC (GTC Val	TAT Tyr 2080	Leu	AAA Lys	GGA Gly	TTG Leu	GAT Asp 2085	Tyr	GAA Glu	AAT Asn	AAA Lys	GAC Asp 2090	Trp	TGT Cys	CCT Pro	6468	
GAA (Glu I	CTA Leu 2095	Glu	GAA Glu	CTT Leu	CAT His	TAC Tyr 2100	Gln	GCA Ala	GCA Ala	TGG Trp	AGG Arg 2105	Asn	ATG Met	CAG Gln	TGG Trp	6516	
GAC (Asp 1 2110	CAT	TGC Cys	ACT Thr	TCC Ser	GTC Val 2115	Ser	AAA Lys	GAA Glu	GTA Val	GAA Glu 2120	Gly	ACC Thr	AGT Ser	TAC Tyr	CAT His 2125	6564	
GAA 7 Glu S	rca Ser	TTG Leu	Tyr	AAT Asn 2130	Ala	CTA Leu	CAA Gln	TCT Ser	CTA Leu 2135	Arg	GAC Asp	AGA Arg	GAA Glu	TTC Phe 2140	Ser	6612	
ACA T	TTT Phe	Tyr	GAA Glu 2145	Ser	CTC Leu	AAA Lys	TAT Tyr	GCC Ala 2150	Arg	GTA Val	AAA Lys	GAA Glu	GTG Val 2155	Glu	GAG Glu	6660	
ATG I	:ys	AAG Lys 2160	CGC . Arg	AGC Ser	CTT Leu	Glu	TCT Ser 2165	GTG Val	TAT Tyr	TCG Ser	CTC Leu	TAT Tyr 2170	Pro	ACA Thr	CTT Leu	6708	

	G GCC ATT GGA GAG n Ala Ile Gly Glu 2180	Leu Glu Ser I	
	C ACA CAT AGA CAA 1 Thr His Arg Gln 2195		
	C CAG CTT CTC AAG r Gln Leu Leu Lys 2210		Glu
Pro Ile Met Al	T CTA CGC ACA GTC a Leu Arg Thr Val 25		
	C TCA CAA AGA GAA n Ser Gln Arg Glu 224	Cys Ile Lys A	
	A CTC TCT ATA CTG u Leu Ser Ile Leu 2260	Ala Arg Thr P	
	G GCA ATA TTT CAA g Ala Ile Phe Gln 2275		
	T GAG TGG CAG CTG r Glu Trp Gln Leu 2290		Ala
Lys Lys Glu Gl	G AGT CTT GCC CTG n Ser Leu Ala Leu 05		
	C AGC TGT GCA GCG a Ser Cys Ala Ala 232	Asn Asn Pro S	
	T CTG AGG GTT TGT s Leu Arg Val Cys 2340	Gly Asn Trp I	
	T GCG GTC ATC ATG O Ala Val Ile Met 2355		
	A AAT TAT GAT GGA y Asn Tyr Asp Gly 2370		Asn
Gly Lys Met Ly	G GCA TTT CTC TCA s Ala Phe Leu Ser 85		
_	T GAA AAC TAC ATG e Glu Asn Tyr Met 240	Lys Ser Ser C	
	G AAA AGA GCC AAA u Lys Arg Ala Lys 2420	Glu Glu Val G	
	G ACA AAC AGA TAC n Thr Asn Arg Tyr 2435		

GAC Glu	TTO Let	G GA' 1 Asi	r GAJ p Glu	A TT/ 1 Let 245	ı Ala	CTC Leu	G CGT	GC# J Ala	A CTO A Let 245	ı Lys	GAC	GA:	CG:	F AAI J Lys 240	A CGC S Arg	7572
TTC Phe	TTA Leu	TG:	r AA/ s Lys 246	s Ala	GTT Val	GAA Glu	AAT Asn	TAT Tyr 247	· Ile	AAC Asn	TGC Cys	TT! Let	Let 247	ı Sei	GGA Gly	7620
GAA Glu	GAA Glu	CAT His 248	s Asp	ATO Met	TGG	GTA Val	TTC Phe 248	Arg	CTI Leu	TGT Cys	TCC Ser	CTC Leu 249	Trp	CTT Let	GAA Glu	7668
AAT Asn	TCT Ser 249	: G13	A GTT Val	TCT Ser	GAA Glu	GTC Val 250	Asn	GGC Gly	ATG Met	ATG Met	AAG Lys 250	Arg	GAC Asp	GGA Gly	ATG Met	7716
Lys 251	11e 0	Pro	Thr	Tyr	Lys 251	Phe 5	Leu	Pro	Leu	Met 252	Tyr 0	Gln	Leu	Ala	GCT Ala 2525	7764
AGA Arg	ATG Met	GGG Gly	ACC Thr	AAG Lys 253	Met	ATG Met	GGA Gly	GGC Gly	CTA Leu 253	Gly	TTT Phe	CAT His	GAA Glu	GTC Val 254	CTC Leu 0	7812
AAT Asn	AAT Asn	CTA Leu	ATC Ile 254	Ser	AGA Arg	ATT	TCA Ser	ATG Met 255	Asp	CAC His	CCC Pro	CAT His	CAC His 255	Thr	TTG Leu	7860
TTT Phe	ATT	ATA Ile 256	CTG Leu 0	GCC Ala	TTA Leu	GCA Ala	AAT Asn 2569	Ala	AAC Asn	AGA Arg	GAT Asp	GAA Glu 257	Phe	CTG Leu	ACT Thr	7908
AAA Lys	CCA Pro 257	Glu	GTA Val	GCC Ala	AGA Arg	AGA Arg 2580	Ser	AGA Arg	ATA Ile	ACT Thr	AAA Lys 2589	Asn	GTG Val	CCT Pro	AAA Lys	7956
CAA Gln 2590	Ser	TCT Ser	CAG Gln	CTT Leu	GAT Asp 2599	Glu	GAT Asp	CGA Arg	ACA Thr	GAG Glu 2600	Ala	GCA Ala	AAT Asn	AGA Arg	ATA Ile 2605	8004
ATA Ile	TGT Cys	ACT Thr	ATC Ile	AGA Arg 261	Ser	AGG Arg	AGA Arg	CCT Pro	CAG Gln 261	Met	GTC Val	AGA Arg	AGT Ser	GTT Val 262	Glu	805,2
GCA Ala	CTT Leu	TGT Cys	GAT Asp 2629	Ala	TAT Tyr	ATT Ile	ATA Ile	TTA Leu 2630	Ala	AAC Asn	TTA Leu	GAT Asp	GCC Ala 2635	Thr	CAG Gln	8100
TGG Trp	AAG Lys	ACT Thr 264	CAG Gln	AGA Arg	AAA Lys	GGC Gly	ATA Ile 2645	Asn	ATT Ile	CCA Pro	GCA Ala	GAC Asp 2650	Gln	CCA Pro	ATT Ile	8148
ACT Thr	AAA Lys 2655	Leu	AAG Lys	AAT Asn	TTA Leu	GAA Glu 2660	Asp	GTT Val	GTT Val	GTC Val	CCT Pro 2665	Thr	ATG Met	GAA Glu	ATT Ile	8196
AAG Lys 2670	vai	GAC Asp	CAC His	ACA Thr	GGA Gly 2675	Glu	TAT Tyr	GGA Gly	AAT Asn	CTG Leu 2680	Val	ACT Thr	ATA Ile	CAG Gln	TCA Ser 2685	8244
TTT Phe	AAA Lys	GCA Ala	GAA Glu	TTT Phe 2690	Arg	TTA Leu .	GCA Ala	Gly	GGT Gly 2695	Val .	TAA Asn	TTA Leu	CCA Pro	AAA Lys 2700	Ile	8292
ATA Ile	GAT Asp	TGT Cys	GTA Val 2705	Gly	TCC Ser	GAT (Asp (Gly .	AAG Lys 2710	GAG Glu	AGG . Arg .	AGA Arg	CAG Gln	CTT Leu 2715	Val	AAG Lys	8340

		TG CAA CAG GTC TTC CAG 8388 et Gln Gln Val Phe Gln 2730
		AA ACT AGG AAG AGG AAA 8436 Lu Thr Arg Lys Arg Lys 2745
Leu Thr Ile Cys Thr T	yr Lys Val Val Pro Le	TC TCT CAG CGA AGT GGT 8484 eu Ser Gln Arg Ser Gly 760 2765
		FT GGT GAA TTT CTT GTT 8532 Le Gly Glu Phe Leu Val 2780
		GG CCA AAT GAT TTC AGT 8580 rg Pro Asn Asp Phe Ser 2795
		TG CAA AAA AAG TCT TTT 8628 al Gln Lys Lys Ser Phe 2810
		GC CAA AAT TTT CAA CCA 8676 vs Gln Asn Phe Gln Pro 2825
Val Phe Arg Tyr Phe C	ys Met Glu Lys Phe Le	rg GAT CCA GCT ATT TGG 8724 eu Asp Pro Ala Ile Trp 340 2845
		TA GCT ACT TCT TCT ATT 8772 al Ala Thr Ser Ser Ile 2860
		AT GTA CAG AAT ATC TTG 8820 is Val Gln Asn Ile Leu 2875
ATA AAT GAG CAG TCA G Ile Asn Glu Gln Ser A 2880	CA GAA CTT GTA CAT AT lla Glu Leu Val His II 2885	TA GAT CTA GGT GTT GCT 8868 le Asp Leu Gly Val Ala 2890
		AG ACA GTT CCT TTT AGA 8916 lu Thr Val Pro Phe Arg 2905
Leu Thr Arg Asp Ile V	Wal Asp Gly Met Gly I	TT ACG GGT GTT GAA GGT 8964 le Thr Gly Val Glu Gly 920 2925
		AA GTG ATG AGA AAC TCT 9012 lu Val Met Arg Asn Ser 2940
		TT CTA TAT GAT CCA CTC 9060 eu Leu Tyr Asp Pro Leu 2955
TTT GAC TGG ACC ATG A Phe Asp Trp Thr Met A 2960	AAT CCT TTG AAA GCT T Asn Pro Leu Lys Ala Lo 2965	TG TAT TTA CAG CAG AGG 9108 eu Tyr Leu Gln Gln Arg 2970
		TG AAT GCA GAT GAC CAA 9156 eu Asn Ala Asp Asp Gln 2985

-193-

GAA Glu 299	Cys	AAA Lys	CGA Arg	AAT Asn	CTC Leu 299	Ser	GAT Asp	ATT Ile	GAC Asp	CAG Gln 3000	Ser	TTC Phe	GAC Asp	AAA Lys	GTA Val 3005	9204
GCT Ala	GAA Glu	CGT Arg	GTC Val	TTA Leu 3010	Met	AGA Arg	CTA Leu	CAA Gln	GAG Glu 3015	Lys	CTG Leu	AAA Lys	GGA Gly	GTG Val 3020	Glu	9252
GAA Glu	GGC Gly	ACT Thr	GTG Val 3029	CTC Leu	AGT Ser	GTT Val	GGT Gly	GGA Gly 3030	Gln	GTG Val	TAA naA	TTG Leu	CTC Leu 3035	Ile	CAG Gln	9300
CAG Gln	GCC Ala	ATA Ile 3040	Asp	CCC Pro	AAA Lys	AAT Asn	CTC Leu 3045	Ser	CGA Arg	CTT Leu	TTC Phe	CCA Pro 3050	Gly	TGG Trp	AAA Lys	9348
	TGG Trp 3055	Val	TGAT	CTTC	'AG T	CATAT	GAAT	T AC	CCT1	TC						9385

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3056 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35: Met Ser Leu Val Leu Asn Asp Leu Leu Ile Cys Cys Arg Gln Leu Glu His Asp Arg Ala Thr Glu Arg Lys Lys Glu Val Glu Lys Phe Lys Arg Leu Ile Arg Asp Pro Glu Thr Ile Lys His Leu Asp Arg His Ser Asp Ser Lys Gln Gly Lys Tyr Leu Asn Trp Asp Ala Val Phe Arg Phe Leu Gln Lys Tyr Ile Gln Lys Glu Thr Glu Cys Leu Arg Ile Ala Lys Pro Asn Val Ser Ala Ser Thr Gln Ala Ser Arg Gln Lys Lys Met Gln Glu Ile Ser Ser Leu Val Lys Tyr Phe Ile Lys Cys Ala Asn Arg Arg Ala 105 Pro Arg Leu Lys Cys Gln Glu Leu Leu Asn Tyr Ile Met Asp Thr Val

Lys Asp Ser Ser Asn Gly Ala Ile Tyr Gly Ala Asp Cys Ser Asn Ile 135

Leu Leu Lys Asp Ile Leu Ser Val Arg Lys Tyr Trp Cys Glu Ile Ser

Gln Gln Gln Trp Leu Glu Leu Phe Ser Val Tyr Phe Arg Leu Tyr Leu 165

Lys Pro Ser Gln Asp Val His Arg Val Leu Val Ala Arg Ile Ile His 180 185

-194-

Ala	Val	Thr 195	Lys	Gly	Cys	Cys	Ser 200	Gln	Thr	Asp	Gly	Leu 205	Asn	Ser	Lys
Phe	Leu 210	Asp	Phe	Phe	Ser	Lys 215	Ala	Ile	Gln	Cys	Ala 220	Arg	Gln	Glu	Lys
Ser 225	Ser	Ser	Gly	Leu	Asn 230	His	Ile	Leu	Ala	Ala 235	Leu	Thr	Ile	Phe	Leu 240
Lys	Thr	Leu	Ala	Val 245	Asn	Phe	Arg	Ile	Arg 250	Val	Cys	Glu	Leu	Gly 255	Asp
Glu	Ile	Leu	Pro 260	Thr	Leu	Leu	Tyr	11e 265	Trp	Thr	Gln	His	Arg 270	Leu	Asn
Asp	Ser	Leu 275	Lys	Glu	Val	Ile	Ile 280	Glu	Leu	Phe	Gln	Leu 285	Gln	Ile	Tyr
Ile	His 290	His	Pro	Lys	Gly	Ala 295	Lys	Thr	Gln	Glu	300 TA8	Gly	Ala	Tyr	Glu
Ser 305	Thr	Lys	Trp	Arg	Ser 310	Ile	Leu	Tyr	Asn	Leu 315	Tyr	Asp	Leu	Leu	Val 320
Asn	Glu	Ile	Ser	His 325	Ile	Gly	Ser	Arg	Gly 330	Lys	Tyr	Ser	Ser	Gly 335	Phe
Arg	Asn	Ile	Ala 340	Val	Lys	Glu	Asn	Leu 3 4 5	Ile	Glu	Leu	Met	Ala 350	qaA	Ile
Cys	His	Gln 355	Val	Phe	Asn	Glu	Asp 360	Thr	Arg	Ser	Leu	Glu 365	Ile	Ser	Gln
Ser	Tyr 370	Thr	Thr	Thr	Gln	Arg 375	Glu	Ser	Ser	Asp	Tyr 380	Ser	Val	Pro	Сув
Lys 385	Arg	Lys	Lys	Ile	Glu 390	Leu	Gly	Trp	Glu	Val 395	Ile	Lys	Asp	His	Leu 400
Gln	Lys	Ser	Gln	Asn 405	Asp	Phe	Asp	Leu	Val 410	Pro	Trp	Leu	Gln	Ile 415	Ala
Thr	Gln	Leu	Ile 420	Ser	Lys	Tyr	Pro	Ala 425	Ser	Leu	Pro	Asn	Cys 430	Glu	Leu
Ser	Pro	Leu 435	Leu	Met	Ile	Leu	Ser 440	Gln	Leu	Leu	Pro	Gln 445	Gln	Arg	His
Gly	Glu 450	Arg	Thr	Pro	Tyr	Val 455	Leu	Arg	Cys	Leu	Thr 460	Glu	Val	Ala	Leu
Cys 465	Gln	Asp	Lys	Arg	Ser 470	Asn	Leu	Glu	Ser	Ser 475	Gln	Lys	Ser	Asp	Leu 480
Leu	Lys	Leu	Trp	Asn 485	Lys	Ile	Trp	Cys	Ile 490	Thr	Phe	Arg	Gly	Ile 495	Ser
Ser	Glu	Gln	11e 500	Gln	Ala	Glu	Asn	Phe 505	Gly	Leu	Leu	Gly	Ala 510	Ile	Ile
Gln	Gly	Ser 515	Leu	Val	Glu	Val	Asp 520	Arg	Glu	Phe	Trp	Lys 525	Leu	Phe	Thr
Gly	Ser 530	Ala	Cys	Arg	Pro	Ser 535	Cys	Pro	Ala	Val	Cys 5 4 0	Cys	Leu	Thr	Leu

Ala 545	Leu	ı Thr	Thr	Sei	550		l Pro	o Gly	/ Ala	a Va:		s Met	E Gly	/ Ile	9 Gl 56
Gln	Asn	Met	Суѕ	Glu 565	ı Val	. Asr	ı Arg	g Ser	Phe 570		Leu	ı Lys	s Glu	Sen 575	
Met	Lys	Trp	Leu 580	Leu)	. Phe	туг	Glm	Leu 585	ı Glu	ı Gly	/ Asp	Let	Glu 590		ı Se
Thr	Glu	Val 595	Pro	Pro	Ile	Leu	His 600		Asr	n Phe	Pro	His 605		ı Val	Le
Glu	Lys 610	Ile	Leu	Val	. Ser	Leu 615	Thr	Met	Lys	as Asr	Cys 620		s Ala	Ala	Me
Asn 625	Phe	Phe	Gln	Ser	Val 630	Pro	Glu	су Су в	Glu	His 635		Glr	Lys	. Asp	640
Glu	Glu	Leu	Ser	Phe 645	Ser	Glu	Val	Glu	Glu 650		Phe	Leu	Gln	Thr 655	
Phe	Asp	Lys	Met 660	Asp	Phe	Leu	Thr	Ile 665		Arg	Glu	Cys	Gly 670		Glu
Lys	His	Gln 675	Ser	Ser	Ile	Gly	Phe 680	Ser	Val	His	Gln	Asn 685		Lys	Glu
Ser	Leu 690	Asp	Arg	Сув	Leu	Leu 695	Gly	Leu	Ser	Glu	Gln 700	Leu	Leu	Asn	Asr
Tyr 705	Ser	Ser	Glu	Ile	Thr 710	Asn	Ser	Glu	Thr	Leu 715	Val	Arg	Cys	Ser	Arc 720
Leu	Leu	Val	Gly	Val 725	Leu	Gly	Сув	Tyr	Cys 730		Met	Gly	Val	Ile 735	Ala
Glu	Glu	Glu	Ala 740	Tyr	Lys	Ser	Glu	Leu 745	Phe	Gln	Lys	Ala	Asn 750	Ser	Leu
Met	Gln	Cys 755	Ala	Gly	Glu	Ser	Ile 760	Thr	Leu	Phe	Lys	Asn 765	Lys	Thr	Asn
Glu	Glu 770	Phe	Arg	Ile	Gly	Ser 775	Leu	Arg	Asn	Met	Met 780	Gln	Leu	Cys	Thr
Arg 785	Cys	Leu	Ser	Asn	Cys 790	Thr	Lys	Lys	Ser	Pro 795	Asn	Lys	Ile	Ala	Ser 800
Gly	Phe	Phe	Leu	Arg 805	Leu	Leu	Thr	Ser	Lys 810	Leu	Met	Asn	Asp	Ile 815	Ala
Asp	Ile	Cys	Lys 820	Ser	Leu	Ala	Ser	Phe 825	Ile	Lys	Lys	Pro	Phe 830	Asp	Arg
Gly	Glu	Val 835	Glu	Ser	Met	Glu	Asp 840	Asp	Thr	Asn	Gly	Asn 845	Leu	Met	Glu
Val	Glu 850	Asp	Gln	Ser	Ser	Met 855	Asn	Leu	Phe	Asn	Asp 860	Tyr	Pro	Asp	Ser
Ser 865	Val	Ser	Asp	Ala	Asn 870	Glu	Pro	Gly	Glu	Ser 875	Gln	Ser	Thr	Ile	Gly 880
Ala	Ile	Asn	Pro	Leu 885	Ala	Glu	Glu	Tyr	ניeL 890	Ser	Lys	Gln	Asp	Leu 895	Leu

WO 97/18323 PCT/US96/19337

-196-

Phe															
	Leu	Asp	Met 900	Leu	Lys	Phe	Leu	Cys 905	Leu	Cys,	Val	Thr	Thr 910	Ala	Gln
Thr	Asn	Thr 915	Val	Ser	Phe	Arg	Ala 920	Ala	Asp	Ile	Arg	Arg 925	Lys	Leu	Leu
Met	Leu 930	Ile	Asp	Ser	Ser	Thr 935	Leu	Glu	Pro	Thr	Lys 940	Ser	Leu	His	Leu
His 945	Met	Tyr	Leu	Met	Leu 950	Leu	Lys	Glu	Leu	Pro 955	Gly	Glu	Glu	Tyr	Pro 960
Leu	Pro	Met	Glu	Asp 965	Val	Leu	Glu	Leu	Leu 970	Lys	Pro	Leu	Ser	Asn 975	Val
Сув	Ser	Leu	Tyr 980	Arg	Arg	qaA	Gln	Asp 985	Val	Cys	Lys	Thr	Ile 990	Leu	Asn
His	Val	Leu 995	His	Val	Val	Lys	Asn 1000		Gly	Gln	Ser	Asn 1005		Asp	Ser
Glu	As n 1010		Arg	Asp	Ala	Gln 1015		Gln	Phe	Leu	Thr 1020		Ile	Gly	Ala
Phe 1025		His	Leu	Thr	Lys 1030		Arg	Lys	Tyr	Ile 1035		Ser	Val	Arg	Met 1040
Ala	Leu	Val	Asn	Cys 1045		Lys	Thr	Leu	Leu 1050		Ala	Asp	Pro	Tyr 1055	
Lys	Trp	Ala	Ile 1060		Asn	Val	Met	Gly 1065		Asp	Phe	Pro	Val 1070	Asn)	Glu
Val	Phe	Thr 1075		Phe	Leu	Ala	_		His	His	Gln		_	Met	Leu
			•				1080	,				1085)		
Ala	Ala 1090	Glu		Ile	Asn	Arg 1095	Leu		Gln	Asp	Thr 1100	Lys		Asp	Ser
	1090 Arg	Glu)	Ser			1095 Leu	Leu	Phe		_	1100 Gln	Lys)	Gly	Asp Ala	
Ser 1109	1090 Ar g	Glu) Leu	Ser	Lys	Ala 1110 Lys	1095 Leu)	Leu Fro	Phe Leu	Lys	Leu 1115 Met	1100 Gln	Lys) Gln	Gly Thr		Phe 1120 His
Ser 1105 Glu	Arg Arg Asn	Glu) Leu Ala	Ser Leu Tyr	Lys Leu 1125 Pro	Ala 1110 Lys	Leu) Ala	Leu Pro Gln	Phe Leu Glu	Lys Gly 1130 Glu	Leu 1115 Met	Gln Gln Arg	Lys Gln Glu	Gly Thr Met	Ala Ser 1135	Phe 1120 His
Ser 1105 Glu Ser	Arg Asn Asn	Glu Leu Ala Glu	Ser Leu Tyr Asn 1140	Lys Leu 1125 Pro	Ala 1110 Lys Glu	Leu Ala Thr	Pro Gln Leu	Phe Leu Glu Asp 1145 Val	Lys Gly 1130 Glu	Leu 1115 Met)	1100 Gln Arg Tyr	Lys Gln Glu Asn	Gly Thr Met Arg 1150	Ala Ser 1135	Phe 1120 His Ser
Ser 1105 Glu Ser Val	Arg Asn Ala Leu	Glu Leu Ala Glu Leu 1155	Ser Leu Tyr Asn 1140	Lys Leu 1125 Pro	Ala 1110 Lys Glu Ile	Leu Ala Thr	Pro Gln Leu Val 1160	Phe Leu Glu Asp 1145	Lys Gly 1130 Glu S	Leu 1115 Met) Ile Ser	Gln Gln Tyr Cys	Lys Gln Glu Asn Ser 1165	Gly Thr Met Arg 1150	Ser 1135	Phe 1120 His Ser
Ser 1105 Glu Ser Val	Arg Asn Ala Leu Lys 1170	Glu Leu Ala Glu Leu 1155	Ser Leu Tyr Asn 1140 Thr	Leu 1125 Pro Leu Leu	Ala 1110 Lys Glu Ile	Leu Ala Thr Ala Ala 1175	Pro Gln Leu Val 1160 Leu	Phe Leu Glu Asp 1145 Val	Lys Gly 1130 Glu Leu Lys	Leu 1115 Met) Ile Ser	Gln Arg Tyr Cys Val 1180	Lys Gln Glu Asn Ser 1165	Gly Thr Met Arg 1150 Pro Glu	Ser 1135 Lys	Phe 1120 His Ser Cys
Ser 1105 Glu Ser Val Glu Leu 1185	Arg Asn Ala Leu Lys 1170	Glu Leu Ala Glu Leu 1155 Gln Pro	Ser Leu Tyr Asn 1140 Thr Ala	Lys Leu 1125 Pro Leu Leu Leu	Ala 1110 Lys Glu Ile Phe Val 1190 Leu	Leu Ala Thr Ala Ala 1175	Pro Gln Leu Val 1160 Leu	Phe Leu Glu Asp 1145 Val Cys	Lys Gly 1130 Glu Leu Lys Leu	Leu 1115 Met Ile Ser Ser Glu 1195	Tyr Cys Val 1186	Lys Gln Glu Asn Ser 1165 Lys Val	Gly Thr Met Arg 1150 Pro Glu Ser	Ser 1135 Lys) Ile Asn	Phe 1120 His Ser Cys Gly Thr 1200
Ser 1105 Glu Ser Val Glu Leu 1185 Phe	Arg Asn Ala Leu Lys 1170 Glu Gly	Glu Leu Ala Glu Leu 1155 Gln Pro	Ser Leu Tyr Asn 1140 Thr Ala His	Lys Leu 1125 Pro Leu Leu Leu Arg 1205	Ala 1110 Lys Glu Ile Phe Val 1190 Leu	Leu Ala Thr Ala Ala 1175 Lys Glu	Pro Gln Leu Val 1160 Leu Lys	Phe Leu Glu Asp 1145 Val Cys Val	Lys Gly 1130 Glu Leu Lys Leu Met 1210	Leu 1115 Met) Ile Ser Ser Glu 1195 Ala	Cys Val 1180 Lys Ser	Lys Cln Glu Asn Ser 1165 Lys Val	Gly Thr Met Arg 1150 Pro Glu Ser Leu	Ala Ser 1135 Lys Ile Asn Glu Asp 1215 Leu	Phe 1120 His Ser Cys Gly Thr 1200

- Arg Ser Cys Tyr Lys Val Leu Ile Pro His Leu Val Ile Arg Ser His 1250 1260
- Phe Asp Glu Val Lys Ser Ile Ala Asn Gln Ile Gln Glu Asp Trp Lys 1265 1270 1275 1280
- Ser Leu Leu Thr Asp Cys Phe Pro Lys Ile Leu Val Asn Ile Leu Pro 1285 1290 1295
- Tyr Phe Ala Tyr Glu Gly Thr Arg Asp Ser Gly Met Ala Gln Gln Arg 1300 1310
- Glu Thr Ala Thr Lys Val Tyr Asp Met Leu Lys Ser Glu Asn Leu Leu 1315 1320 1325
- Gly Lys Gln Ile Asp His Leu Phe Ile Ser Asn Leu Pro Glu Ile Val 1330 1335 1340
- Val Glu Leu Leu Met Thr Leu His Glu Pro Ala Asn Ser Ser Ala Ser 1345 1350 1355 1360
- Gln Ser Thr Asp Leu Cys Asp Phe Ser Gly Asp Leu Asp Pro Ala Pro 1365 1370 1375
- Asn Pro Pro His Phe Pro Ser His Val Ile Lys Ala Thr Phe Ala Tyr 1380 1385 1390
- Ile Ser Asn Cys His Lys Thr Lys Leu Lys Ser Ile Leu Glu Ile Leu 1395 1400 1405
- Ser Lys Ser Pro Asp Ser Tyr Gln Lys Ile Leu Leu Ala Ile Cys Glu 1410 1415 1420
- Gln Ala Ala Glu Thr Asn Asn Val Tyr Lys Lys His Arg Ile Leu Lys 1425 1430 1435 1440
- Ile Tyr His Leu Phe Val Ser Leu Leu Leu Lys Asp Ile Lys Ser Gly 1445 1450 1455
- Leu Gly Gly Ala Trp Ala Phe Val Leu Arg Asp Val Ile Tyr Thr Leu 1460 1465 1470
- Ile His Tyr Ile Asn Gln Arg Pro Ser Cys Ile Met Asp Val Ser Leu 1475 1480 1485
- Arg Ser Phe Ser Leu Cys Cys Asp Leu Leu Ser Gln Val Cys Gln Thr 1490 1495 1500
- Ala Val Thr Tyr Cys Lys Asp Ala Leu Glu Asn His Leu His Val Ile 1505 1510 1515 1520
- Val Gly Thr Leu Ile Pro Leu Val Tyr Glu Gln Val Glu Val Gln Lys 1525 1530 1535
- Gln Val Leu Asp Leu Leu Lys Tyr Leu Val Ile Asp Asn Lys Asp Asn 1540 1545 1550
- Glu Asn Leu Tyr Ile Thr Ile Lys Leu Leu Asp Pro Phe Pro Asp His 1555 1560 1565
- Val Val Phe Lys Asp Leu Arg Ile Thr Gln Gln Lys Ile Lys Tyr Ser 1570 1575 1580
- Arg Gly Pro Phe Ser Leu Leu Glu Glu Ile Asn His Phe Leu Ser Val 1585 1590 1595 1600

- Ser Val Tyr Asp Ala Leu Pro Leu Thr Arg Leu Glu Gly Leu Lys Asp 1605 1610 1615
- Leu Arg Arg Gln Leu Glu Leu His Lys Asp Gln Met Val Asp Ile Met 1620 1630
- Arg Ala Ser Gln Asp Asn Pro Gln Asp Gly Ile Met Val Lys Leu Val
- Val Asn Leu Leu Gln Leu Ser Lys Met Ala Ile Asn His Thr Gly Glu 1650 1660
- Lys Glu Val Leu Glu Ala Val Gly Ser Cys Leu Gly Glu Val Gly Pro 1665 1670 1675 1680
- Ile Asp Phe Ser Thr Ile Ala Ile Gln His Ser Lys Asp Ala Ser Tyr 1685 1690 1695
- Thr Lys Ala Leu Lys Leu Phe Glu Asp Lys Glu Leu Gln Trp Thr Phe 1700 1705 1710
- Ile Met Leu Thr Tyr Leu Asn Asn Thr Leu Val Glu Asp Cys Val Lys 1715 1720 1725
- Val Arg Ser Ala Ala Val Thr Cys Leu Lys Asn Ile Leu Ala Thr Lys 1730 1735 1740
- Thr Gly His Ser Phe Trp Glu Ile Tyr Lys Met Thr Thr Asp Pro Met 1745 1750 1755 1760
- Leu Ala Tyr Leu Gln Pro Phe Arg Thr Ser Arg Lys Lys Phe Leu Glu 1765 1770 1775
- Val Pro Arg Phe Asp Lys Glu Asn Pro Phe Glu Gly Leu Asp Asp Ile 1780 1785 1790
- Asn Leu Trp Ile Pro Leu Ser Glu Asn His Asp Ile Trp Ile Lys Thr 1795 1800 1805
- Leu Thr Cys Ala Phe Leu Asp Ser Gly Gly Thr Lys Cys Glu Ile Leu 1810 1815 1820
- Gln Leu Leu Lys Pro Met Cys Glu Val Lys Thr Asp Phe Cys Gln Thr 1825 1830 1835 1840
- Val Leu Pro Tyr Leu Ile His Asp Ile Leu Leu Gln Asp Thr Asn Glu 1845 1850 1855
- Ser Trp Arg Asn Leu Leu Ser Thr His Val Gln Gly Phe Phe Thr Ser 1860 1865 1870
- Cys Leu Arg His Phe Ser Gln Thr Ser Arg Ser Thr Thr Pro Ala Asn 1875 1880 1885
- Leu Asp Ser Glu Ser Glu His Phe Phe Arg Cys Cys Leu Asp Lys Lys 1890 1895 1900
- Ser Gln Arg Thr Met Leu Ala Val Val Asp Tyr Met Arg Arg Gln Lys 1905 1910 1915 1920
- Arg Pro Ser Ser Gly Thr Ile Phe Asn Asp Ala Phe Trp Leu Asp Leu 1925 1930 1935
- Asn Tyr Leu Glu Val Ala Lys Val Ala Gln Ser Cys Ala Ala His Phe 1940 1945 1950

- Thr Ala Leu Leu Tyr Ala Glu Ile Tyr Ala Asp Lys Lys Ser Met Asp 1955 1960 1965
- Asp Gln Glu Lys Arg Ser Leu Ala Phe Glu Glu Gly Ser Gln Ser Thr 1970 1975 1980
- Thr Ile Ser Ser Leu Ser Glu Lys Ser Lys Glu Glu Thr Gly Ile Ser 1985 1990 1995 2000
- Leu Gln Asp Leu Leu Glu Ile Tyr Arg Ser Ile Gly Glu Pro Asp 2005 2010 2015
- Ser Leu Tyr Gly Cys Gly Gly Gly Lys Met Leu Gln Pro Ile Thr Arg 2020 2025 2030
- Leu Arg Thr Tyr Glu His Glu Ala Met Trp Gly Lys Ala Leu Val Thr 2035 2040 2045
- Tyr Asp Leu Glu Thr Ala Ile Pro Ser Ser Thr Arg Gln Ala Gly Ile 2050 2055 2060
- Ile Gln Ala Leu Gln Asn Leu Gly Leu Cys His Ile Leu Ser Val Tyr 2065 2070 2075 2080
- Leu Lys Gly Leu Asp Tyr Glu Asn Lys Asp Trp Cys Pro Glu Leu Glu 2085 2090 2095
- Glu Leu His Tyr Gln Ala Ala Trp Arg Asn Met Gln Trp Asp His Cys 2100 2105 2110
- Thr Ser Val Ser Lys Glu Val Glu Gly Thr Ser Tyr His Glu Ser Leu 2115 2120 2125
- Tyr Asn Ala Leu Gln Ser Leu Arg Asp Arg Glu Phe Ser Thr Phe Tyr 2130 2135 2140
- Glu Ser Leu Lys Tyr Ala Arg Val Lys Glu Val Glu Glu Met Cys Lys 2145 2150 2155 2160
- Arg Ser Leu Glu Ser Val Tyr Ser Leu Tyr Pro Thr Leu Ser Arg Leu 2165 2170 2175
- Gln Ala Ile Gly Glu Leu Glu Ser Ile Gly Glu Leu Phe Ser Arg Ser 2180 2185 2190
- Val Thr His Arg Gln Leu Ser Glu Val Tyr Ile Lys Trp Gln Lys His 2195 2200 2205
- Ser Gln Leu Leu Lys Asp Ser Asp Phe Ser Phe Gln Glu Pro Ile Met 2210 2225 2220
- Ala Leu Arg Thr Val Ile Leu Glu Ile Leu Met Glu Lys Glu Met Asp 2225 2230 2235 2240
- Asn Ser Gln Arg Glu Cys Ile Lys Asp Ile Leu Thr Lys His Leu Val 2245 2250 2255
- Glu Leu Ser Ile Leu Ala Arg Thr Phe Lys Asn Thr Gln Leu Pro Glu 2260 2265 2270
- Arg Ala Ile Phe Gln Ile Lys Gln Tyr Asn Ser Val Ser Cys Gly Val 2275 2280 2285
- Ser Glu Trp Gln Leu Glu Glu Ala Gln Val Phe Trp Ala Lys Lys Glu 2290 2295 2300

WO 97/18323 PCT/US96/19337

-200-

									200	-					
Gln 2305		Leu	Ala	Leu	Ser 2310		Leu	Lys	Gln	Met 2319		Lys	Lys	Leu	Asp 2320
Ala	Ser	Cys	Ala	Ala 2325		Asn	Pro	Ser	Leu 2330		Leu	Thr	Tyr	Thr 2335	
Cys	Leu	Arg	Val 2340		Gly	Asn	Trp	Leu 2345		Glu	Thr	Cys	Leu 2350		Asn
Pro	Ala	Val 2355	Ile	Met	Gln	Thr	Tyr 2360		Glu	Lys	Ala	Val 2365		Val	Ala
Gly	Asn 2370	-	Asp	Gly	Glu	Ser 23 7 5		Asp	Glu	Leu	Arg 2380		Gly	Lys	Met
Lys 2385		Phe	Leu	Ser	Leu 2390		Arg	Phe	Ser	Asp 2395		Gln	Tyr	Gln	Arg 2400
Ile	Glu	Asn	Tyr	Met 2405		Ser	Ser	Glu	Phe 2410		Asn	Lys	Gln	Ala 2415	
Leu	Lys	Arg	Ala 2420	_	Glu	Glu	Val	Gly 2425		Leu	Arg	Glu	His 2430	_	Ile
Gln	Thr	Asn 2435	Arg	Tyr	Thr	Val	Lys 2440		Gln	Arg	Glu	Leu 2445		Leu	Asp
Glu	Leu 2450		Leu	Arg	Ala	Leu 2455		Glu	Asp	Arg	Lys 2460		Phe	Leu	Cys
Lys 2465		Val	Glu	Asn	Tyr 2470		Asn	Cys	Leu	Leu 2475		Gly	Glu	Glu	Нів 2480
Asp	Met	Trp	Val	Phe 2485		Leu	Cys	Ser	Leu 2490		Leu	Glu	Asn	Ser 2495	
Val	Ser	Glu	Val 2500		Gly	Met	Met	Lys 2505		Asp	Gly	Met	Lys 2510		Pro
Thr	Tyr	Lys 2515	Phe	Leu	Pro	Leu	Met 2520	_	Gln	Leu	Ala	Ala 2525	_	Met	Gly
Thr	Lys 2530		Met	Gly	Gly	Leu 2535		Phe	His	Glu	Val 2540		Asn	Asn	Leu
Ile 2545		Arg	Ile	Ser	Met 2550		His	Pro	His	His 2555		Leu	Phe	Ile	Ile 2560
Leu	Ala	Leu	Ala	Asn 2565		Asn	Arg	Asp	Glu 2 5 70		Leu	Thr	Lys	Pro 2575	
Val	Ala	Arg	Arg 2580		Arg	Ile	Thr	Lys 2585		Val	Pro	Lys	Gln 2590		Ser
Gln	Leu	Asp 2595	Glu	Asp	Arg	Thr	Glu 2600		Ala	Asn	Arg	11e 260		Cys	Thr
Ile	Arg 261		Arg	Arg	Pro	Gln 2619		Val	Arg	Ser	Val 2620		Ala	Leu	Cys
Asp 2629		Tyr	Ile	Ile	Leu 2630		Asn	Leu	Asp	Ala 2635		Gln	Trp	Lys	Thr 2640

Gln Arg Lys Gly Ile Asn Ile Pro Ala Asp Gln Pro Ile Thr Lys Leu 2645 2650 2655

- Lys Asn Leu Glu Asp Val Val Val Pro Thr Met Glu Ile Lys Val Asp 2660 2665 2670
- His Thr Gly Glu Tyr Gly Asn Leu Val Thr Ile Gln Ser Phe Lys Ala 2675 2680 2685
- Glu Phe Arg Leu Ala Gly Gly Val Asn Leu Pro Lys Ile Ile Asp Cys 2690 2695 2700
- Val Gly Ser Asp Gly Lys Glu Arg Arg Gln Leu Val Lys Gly Arg Asp 2705 2710 2715 2720
- Asp Leu Arg Gln Asp Ala Val Met Gln Gln Val Phe Gln Met Cys Asn 2725 2730 2735
- Thr Leu Leu Gln Arg Asn Thr Glu Thr Arg Lys Arg Lys Leu Thr Ile 2740 2745 2750
- Cys Thr Tyr Lys Val Val Pro Leu Ser Gln Arg Ser Gly Val Leu Glu 2755 2760 2765
- Trp Cys Thr Gly Thr Val Pro Ile Gly Glu Phe Leu Val Asn Asn Glu 2770 2775 2780
- Asp Gly Ala His Lys Arg Tyr Arg Pro Asn Asp Phe Ser Ala Phe Gln 2785 2790 2795 2800
- Cys Gln Lys Lys Met Met Glu Val Gln Lys Lys Ser Phe Glu Glu Lys 2805 2810 2815
- Tyr Glu Val Phe Met Asp Val Cys Gln Asn Phe Gln Pro Val Phe Arg 2820 2825 2830
- Tyr Phe Cys Met Glu Lys Phe Leu Asp Pro Ala Ile Trp Phe Glu Lys 2835 2840 2845
- Arg Leu Ala Tyr Thr Arg Ser Val Ala Thr Ser Ser Ile Val Gly Tyr 2850 2855 2860 .
- Ile Leu Gly Leu Gly Asp Arg His Val Gln Asn Ile Leu Ile Asn Glu 2865 2870 2880
- Gln Ser Ala Glu Leu Val His Ile Asp Leu Gly Val Ala Phe Glu Gln 2885 2890 2895
- Gly Lys Ile Leu Pro Thr Pro Glu Thr Val Pro Phe Arg Leu Thr Arg 2900 2905 2910
- Asp Ile Val Asp Gly Met Gly Ile Thr Gly Val Glu Gly Val Phe Arg 2915 2920 2925
- Arg Cys Cys Glu Lys Thr Met Glu Val Met Arg Asn Ser Gln Glu Thr 2930 2935 2940
- Leu Leu Thr Ile Val Glu Val Leu Leu Tyr Asp Pro Leu Phe Asp Trp
 2945 2950 2955 2960
- Thr Met Asn Pro Leu Lys Ala Leu Tyr Leu Gln Gln Arg Pro Glu Asp 2965 2970 2975
- Glu Thr Glu Leu His Pro Thr Leu Asn Ala Asp Asp Gln Glu Cys Lys 2980 2985 2990
- Arg Asn Leu Ser Asp Ile Asp Gln Ser Phe Asp Lys Val Ala Glu Arg 2995 3000 3005

-202-

Val Leu Met Arg Leu Gln Glu Lys Leu Lys Gly Val Glu Glu Gly Thr 3015

Val Leu Ser Val Gly Gly Gln Val Asn Leu Leu Ile Gln Gln Ala Ile 3025

Asp Pro Lys Asn Leu Ser Arg Leu Phe Pro Gly Trp Lys Ala Trp Val 3045 3050

- (2) INFORMATION FOR SEQ ID NO:36:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met Ser Gly Gly Ser Ser Cys Gln Thr Pro Ser Arg Ala Ile Pro Ala

Thr Arg Arg

- (2) INFORMATION FOR SEQ ID NO:37:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 amino acids
 - (B) TYPE: amino acid (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Gly Asp Tyr Ser Thr Thr Pro Gly Gly Thr Leu Phe Ser Thr Thr Pro 5 10

Gly Gly Thr Arg Arg 20

- (2) INFORMATION FOR SEQ ID NO:38:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Glu Cys Arg Asn Ser Pro Val Thr Lys Thr Arg Arg

-203-

- (2) INFORMATION FOR SEQ ID NO:39:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
 - Gly Val Thr Ser Pro Ser Ser Asp Glu Pro Arg Arg
- (2) INFORMATION FOR SEQ ID NO:40:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Glu Ala Ser Gln Ser His Leu Arg Arg

- (2) INFORMATION FOR SEQ ID NO:41:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid(D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Arg Arg Asn Ser Pro Glu Asp Lys Arg Ala Gly Gly

- (2) INFORMATION FOR SEQ ID NO:42:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:
 - Gly Glu Glu Ser Gln Phe Glu Met Asp Ile Arg Arg



- 204 -

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism ref	erred to in the description -6
B. IDENTIFICATION OF DEPOSIT	Further deposits are identified on an additional sheet
Name of depositary institution	
American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, MD 20852 US	
Date of deposit	Accession Number S
7 November 1996	HB 12233 and HB 12234
C. ADDITIONAL INDICATIONS (leave blank if not applicable	(c) This information is continued on an additional sheet
"In respect of those designations a sample of the deposited microorganism publication of the mention of the grandate on which the application has been be withdrawn, only by the issue of suct the person requesting the sample (Rule	t of the European patent or until the refused or withdrawn or is deemed to have an expert nominated by
D. DESIGNATED STATES FOR WHICH INDICATION	NS ARE MADE (if the indications are not for all designated States)
EP	
E. SEPARATE FURNISHING OF INDICATIONS (leave	e blank if not applicable)
The indications listed below will be submitted to the International Number of Deposit*)	Bureau later (specify the general nature of the indications e.g., "Accession
For receiving Office use only	For International Bureau use only
This sheet was received with the international application	This sheet was received by the International Bureau on:
Authorized of Greek Balls (703) 305-3880	Authorized officer

-205-

CLAIMS

We claim:

- 1. A purified and isolated polynucleotide comprising a polynucleotide encoding the PIK-related kinase MCCS1 α amino acid sequence set out in SEQ ID NO: 31.
- 2. A purified and isolated polynucleotide comprising a polynucleotide encoding the PIK-related kinase MCCS1 β amino acid sequence set out in SEQ ID NO: 33.
 - 3. The polynucleotide of claim 1 or 2 which is a DNA.
 - 4. The DNA of claim 3 which is a cDNA.
- 5. A MCCS1 α cDNA consisting of the DNA sequence set out in SEQ ID NO: 30.
- 6. A MCCS1 β DNA consisting of the DNA sequence set out in SEQ ID NO: 32.
 - 7. The DNA of claim 3 which is a genomic DNA.
 - 8. An RNA transcript of the DNA of claim 3.
- 9. The DNA of claim 3 which is a wholly or partially chemically synthesized DNA.

- 10. A DNA comprising a DNA encoding a full length mammalian MCCS1 kinase selected from the group consisting of:
- a) a DNA which hybridizes under stringent conditions to the non-coding strand of the DNA of SEQ ID NO: 30;
- b) a DNA which hybridizes under stringent conditions to the non-coding strand of the DNA of SEQ ID NO: 3; and
- c) a DNA which hybridizes under stringent conditions to the non-coding strand of the DNA of SEQ ID NO: 32.
 - 11. A vector comprising a DNA according to claim 3 or 10.
- 12. The vector of claim 11 wherein said DNA is operatively linked to an expression control DNA sequence.
- 13. A host cell stably transformed or transfected with a DNA according to claim 3 or 10 in a manner allowing the expression in said host cell of the MCCS1 kinase.
- 14. A method for producing the PIK-related kinase MCCS1, said method comprising growing a host cell according to claim 11 in a suitable nutrient medium and isolating the MCCS1 kinase from said cell or the medium of its growth.
- 15. A purified and isolated polypeptide comprising the PIK-related kinase MCCS1α amino acid sequence consisting of SEQ ID NO: 31.
- 16. A purified and isolated polypeptide comprising the PIK-related kinase MCCS1β amino acid sequence consisting of SEQ ID NO: 33.
- 17. A polypeptide or peptide capable of specifically binding to PIK-related kinase MCCS1.
 - 18. An antibody product according to claim 17.

- 19. A monoclonal antibody according to claim 18.
- 20. A hybridoma cell line producing a monoclonal antibody according to claim 19.
- 21. An assay for identifying modulators of MCCS1 kinase activity comprising the steps of:
- a) incubating a MCCS1 kinase preparation in kinase buffer with gamma-32P-ATP and an exogenous kinase substrate in the presence and absence of a test compound, and
- b) measuring the moles of phosphate transferred to said substrate; wherein an increase in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an activator of said MCCS1 kinase and a decrease in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an inhibitor of said MCCS1 kinase.
 - 22. The hybridoma cell line 224C.
 - 23. The hybridoma cell line 224F.

- 24. A method of identifying a compound that inhibits MCCS1 comprising the steps of:
 - a) expressing MCCS1 in a genetically altered cell, thereby decreasing the sensitivity of the cell to DNA damage, said sensitivity being associated with the genetic alteration;
 - b) exposing the genetically altered cell of step (a) to DNA damaging treatment in the presence and absence of a test modulator compound;
 - c) measuring the sensitivity of the cell to DNA damage; and
 - d) identifying a test compound that restores the sensitivity of the cell to DNA damage as an inhibitor of MCCS1 activity.
- 25. A method of identifying a compound that inhibits ATM comprising the steps of:
 - a) expressing ATM in a genetically altered cell, thereby decreasing the sensitivity of the cell to DNA damage, said sensitivity being associated with the genetic alteration;
 - b) exposing the genetically altered cell of step (a) to DNA damaging treatment in the presence and absence of a test modulator compound;
 - c) measuring the sensitivity of the cell to DNA damage; and
 - d) identifying a test compound that restores the sensitivity of the cell to DNA damage as an inhibitor of ATM activity.



- 26. An assay for identifying modulators of ATM kinase activity comprising the steps of:
- a) incubating a ATM kinase preparation in kinase buffer with gamma³²P-ATP and an exogenous kinase substrate in the presence and absence of a test
 compound, and
- b) measuring the moles of phosphate transferred to said substrate; wherein an increase in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an activator of said ATM kinase and a decrease in the moles of ³²P-phosphate transferred to said substrate in presence of said test compound compared to the moles of ³²P-phosphate transferred to said substrate in the absence of said test compound indicates that said test compound is an inhibitor of said ATM kinase.

PCT





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International F	Patent Classification 6:
C12N 9/12, 5/12, C12O	15/63, C07K 16/40,C12N 1/48, C12N 15/11

A3

(11) International Publication Number:

WO 97/18323

(43) Internati nal Publication Date:

22 May 1997 (22.05.97)

(21) International Application Number:

PCT/US96/19337

(22) International Filing Date:

18 November 1996 (18.11.96)

(30) Priority Data:

08/558,666 16 November 1995 (16.11.95) US 08/007,312 27 February 1996 (27.02.96) US 08/725,304 21 October 1996 (21.10.96) US

(71) Applicant (for all designated States except US): ICOS CORPORATION [US/US]; 22021 20th Avenue, S.E., Bothell, WA 98021 (US).

(72) Inventors: and

- (75) Inventors/Applicants (for US only): HOEKSTRA, Merl, F. [US/US]; 10321 216th Street, S.E., Snohomish, WA 98290 (US). HOLTZMAN, Doug, A. [US/US]; 4308 4th Avenue, N.E., Seattle, WA 98105 (US). KEEGAN, Kathleen, S. [US/US]; 5812 Mercer Way, Mercer Island, WA 98040 (US).
- (74) Agent: NOLAND, Greta, E.; Marshall, O'Toole, Gerstein, Murray & Borun, 6300 Sears Tower, 233 South Wacker Drive, Chicago, IL 60606-6402 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
9 October 1997 (09.10.97)

(54) Title: CELL-CYCLE CHECKPOINT PHOSPHATIDYLINOSITOL- (PIK-) RELATED KINASES, GENES CODING THEREFOR AND METHODS FOR ASSAYING AND MODULATING ENZYMATIC ACTIVITY

(57) Abstract

The present invention generally relates to genes encoding cell cycle checkpoint phosphatidylinositol kinase (PIK)-related proteins essential to DNA damage responses in cells. These PIK-related kinases are required in regulatory pathways that arrest the cell cycle foll wing DNA damage to allow DNA repair prior to mitosis or initiation of DNA replication. More particularly, the invention provides a novel human cell cycle checkpoint PIK-related kinase, MCCS1, and polynucleotide sequences encoding the MCSS1. Assays for identifying modulators of MCCS1 useful as, for example, chemotherapy and radiation adjuvants, are also provided by the invention. Further, assays for identifying modulators of the cell cycle checkpoint phosphatidylinositol kinase (PIK)-related protein identified as ATM are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic	SD	Sudan
CF	Central African Republic		of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	ΤT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	ÜA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

INTERNATIONAL SEARCH REPORT

Internal Application No PCT/US 96/19337

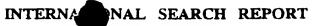
A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N9/12 C12N15/63 C12Q1/48 CO7K16/40 C12N5/12 C12N15/11 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Category ' Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. P,X WO 97 09433 A (MEDICAL RESEARCH COUNCIL) 2-4, 13 March 1997 9-14, 16-18, 21,24 see SEQ ID No. 1 and 2; pages 16-22; page 33, lines 28-33; page 34, lines 1-12 P.X PROCEEDINGS OF THE NATIONAL ACADEMY OF 1,3,4, SCIENCES, 9-11,15, vol. 93, April 1996, pages 2850-2855, XP002023632 CIMPRICH, K.A. ET AL .: "CDNA cloning and gene mapping of a candidate human cell cycle checkpoint protein" see Figure 2 -/--X Further documents are listed in the continuation of box C. Patent family members are listed in annex. Special categories of cited documents: T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the 'A' document defining the general state of the art which is not considered to be of particular relevance invention ·Eearlier document but published on or after the international 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to filing date document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) involve an inventive step when the document is taken alone 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such docudocument referring to an oral disclosure, use, exhibition or other means ments, such combination being obvious to a person skilled document published prior to the international filing date but later than the priority date claimed '&' document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 2 6, 08, 97 1 August 1997 Name and mailing address of the ISA Authorized officer European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Risswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo ni, Alt, G Fax: (+31-70) 340-3016

Form PCT.ISA 210 (second sheet) (July 1992)

1

C/Control	aton) DOCUMENTS CONSIDERED TO BE RELEVANT	PCT/US 9	J/ T2J2/
Category *			Relevant to claim No.
A	GENE, vol. 119, 1992, pages 83-89, XP002023633 SEATON, B.L. ET AL.: "Isolation and characterization of the Schizosaccharomyces pombe rad3 gene, involved in the DNA damage and DNA synthesis checkpoints" cited in the application see Figure 5; page 84, right-hand column		1
P,A	EMBO JOURNAL, vol. 15, no. 23, 1996, pages 6641-6651, XP002023634 BENTLEY, N.J. ET AL.: "The Schizosaccharomyces pombe rad3 checkpoint gene" see the whole document		1
A	SCIENCE, vol. 268, 23 June 1995, pages 1749-1753, XP002036686 SAVITSKY, K. ET AL.: "A single ataxia telangiectasia gene with a product similar to PI-3 kinase" see the whole document		25,26
Т	CURRENT OPINION GENET. DEV., vol. 7, no. 2, 1997, pages 170-175, XP002036687 HOEKSTRA, M.F.: "Responses to DNA damage and regulation of cell cycle checkpoints by the ATM protein kinase family" see the whole document		1
			i a
	*		

1



l Application No

	information on patent family memb	PCT	7US 96/19337	
Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
WO 9709433 A	13-03-97	AU 6884696 A	27-03-97	
	*	*************		
	•			
	•			
*				
			•	
*				
•				

. . * ..